Spatial epidemiology of a highly transmissible disease in urban neighbourhoods: Using COVID-19 outbreaks in Toronto as a case study

by

Nushrat Nazia

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Public Health Sciences

Waterloo, Ontario, Canada, 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Dr. Ying MacNab Associate Professor School of Population and Public Health The University of British Columbia
Supervisor	Dr. Jane Law Associate Professor School of Public Health Sciences University of Waterloo
Internal Members	Dr. Zahid Ahmad Butt Assistant Professor School of Public Health Sciences University of Waterloo
	Dr. Charmaine Dean Professor Department of Statistics and Actuarial Science University of Waterloo
Internal-external Member	Dr. Susan Elliott Professor

Department of Geography and Environmental Management
University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The three manuscripts (Chapters 3, 4 and 5) presented in this thesis are the work of Nushrat Nazia in collaboration with her co-authors and committee members. These manuscripts have either been submitted or published in peer-review journals,
Exceptions to sole authorship include:

Manuscript 1 (Chapter 3): Nazia, N., Law, J., & Butt, Z.A. (2022). Spatiotemporal clusters and the socioeconomic determinants of COVID-19 in Toronto neighbourhoods, Canada. *Spatial and Spatio-Temporal Epidemiology*, 43, 100534. https://doi.org/10.1016/j.sste.2022.100534.

Manuscript 2 (Chapter 4): Nazia, N., Law, J., & Butt, Z.A. (2022). Identifying spatiotemporal patterns of COVID-19 transmissions and the drivers of the patterns in Toronto: a Bayesian hierarchical spatiotemporal modelling. *Scientific Reports* 12, 9369. https://doi.org/10.1038/s41598-022-13403-x.

Manuscript 3 (Chapter 5): Nazia, N., Law, J., & Butt, Z.A. Modelling the spatiotemporal spread of COVID-19 outbreaks and prioritization of the risk areas in Toronto, Canada. Submitted (under review) to *Health and Place* journal.

I declare that as the lead author of these three chapters, I was responsible for conceptualizing the study design, data collection and analysis. I was also responsible for preparing the draft of articles, submission to peer-reviewed journals as well as addressing comments from reviewers. My coauthors, Dr. Jane Law and Dr. Zahid Ahmad Butt, provided methodological guidance, critical feedback and reflections on drafts and revisions of manuscripts. As the principal supervisor, Dr. Jane Law provided significant research direction and editorial assistance throughout.

Abstract

The emergence of infectious diseases in an urban area involves a complex interaction between the socioecological processes in the neighbourhood and urbanization. As a result, such an urban environment can be the incubator of new epidemics and spread diseases more rapidly in densely populated areas than elsewhere. Most recently, the Coronavirus-19 (COVID-19) pandemic has brought unprecedented challenges around the world. Toronto, the capital city of Ontario, Canada, has been severely impacted by COVID-19. Understanding the spatiotemporal patterns and the key drivers of such patterns is imperative for designing and implementing an effective public health program to control the spread of the pandemic. This dissertation was designed to contribute to the global research effort on the COVID-19 pandemic by conducting spatial epidemiological studies to enhance our understanding of the disease's epidemiology in a spatial context to guide enhancing the public health strategies in controlling the disease.

Comprised of three original research manuscripts, this dissertation focuses on the spatial epidemiology of COVID-19 at a neighbourhood scale in Toronto. Each manuscript makes scientific contributions and enhances our knowledge of how interactions between different socioecological processes in the neighbourhood and urbanization can influence spatial spread and patterns of COVID-19 in Toronto with the application of novel and advanced methodological approaches. The findings of the outcomes of the analyses are intended to contribute to the public health policy that informs neighbourhood-based disease intervention initiatives by the public health authorities, local government, and policymakers.

The first manuscript analyzes the globally and locally variable socioeconomic drivers of COVID-19 incidence and examines how these relationships vary across different neighbourhoods. In the global model, lower levels of education and the percentage of immigrants were found to have a positive association with increased risk for COVID-19. This study provides the methodological framework for identifying the local variations in the association between risk for COVID-19 and socioeconomic factors in an urban environment by applying a local multiscale geographically weighted regression (MGWR) modelling approach. The MGWR

model is an improvement over the methods used in earlier studies of COVID-19 in identifying local variations of COVID-19 by incorporating a correction factor for the multiple testing problem in the geographically weighted regression models.

The second manuscript quantifies the associations between COVID-19 cases and urban socioeconomic and land surface temperature (LST) at the neighbourhood scale in Toronto. Four spatiotemporal Bayesian hierarchical models with spatial, temporal, and varying space-time interaction terms are compared. The results of this study identified the seasonal trends of COVID-19 risk, where the spatiotemporal trends show increasing, decreasing, or stable patterns, and identified area-specific spatial risk for targeted interventions. Educational level and high land surface temperature are shown to have a positive association with the risk for COVID-19. In this study, high spatial and temporal resolution satellite images were used to extract LST, and atmospheric corrections methods were applied to these images by adopting a land surface emissivity (LSE) model, which provided a high estimation accuracy. The methodological approach of this work will help researchers understand how to acquire long time-series data of LST at a spatial scale from satellite images, develop methodological approaches for atmospheric correction and create the environmental data with a high estimation accuracy to fit into modelling disease. Applying to policy, the findings of this study can inform the design and implementation of urban planning strategies and programs to control disease risks.

The third manuscript developed a novel approach for visualization of the spread of infectious disease outbreaks by incorporating neighbourhood networks and the time-series data of the disease at the neighbourhood level. The findings of the model provide an understanding of the direction and magnitude of spatial risk for the outbreak and guide for the importance of early intervention in order to stop the spread of the outbreak. The manuscript also identified hotspots using incidence rate and disease persistence, the findings of which may inform public health planners to develop priority-based intervention plans in a resource constraint situation.

Multiple spatial methods were used in the three manuscripts of this thesis to understand the spatial patterns of the risk for COVID-19 in Toronto. Cohen's Kappa coefficient was computed to evaluate the agreement in the spatial patterns of risk for COVID-19 between the methods. The results indicated 62% agreement between Study 1 and Study 2, 61% agreement between Study 2 and Study 3, and 46% agreement between Study 1 and Study 3. Such differences are expected as the mathematics behind computing the spatial patterns of risk differ by each method. However, since the Bayesian approach provides a better computational algorithm and accounts for the uncertainties related to the data, the process and model parameters; the findings provided by this method could be considered robust in identifying the spatial risk of COVID-19 in Toronto.

Overall, this dissertation advances spatial epidemiological research focusing on the interaction between socioecological processes in the neighbourhood and urbanization and its impact on the spread of the COVID-19 outbreak in Toronto by i) examining the relationship of the disease with socioeconomic and environmental factors (such as land surface temperature), ii) identifying the sources and directions of spatial spread of the COVID-19 outbreaks, iii) identifying spatial and temporal trends of the disease across neighbourhoods, and iv) identifying high-risk areas using multiple methods. The findings of this dissertation could be used to design and implement infectious disease prevention initiatives in any setting. Finally, the methodological framework used in this research could be used in any other disease outbreak where local-level data are available.

Acknowledgements

To Dr. Jane Law: I would like to express my sincere gratitude for your vision, guidance, encouragement, patience and continued support. Thank you for believing in me and providing me with the opportunity to learn from you. I always appreciate your constructive and critical feedback that has helped me improve my work. You have helped me grow significantly as an independent researcher and as a person.

To Dr. Zahid Butt: Thank you for your generosity, patience and advice in supporting my Ph.D. journey for the past few years. Your support for my academic work, as well as a mentor, has helped me navigate through many challenging times. I appreciate the knowledge you imparted to me, and I am grateful for your wonderful guidance and support

To Dr. Charmaine Dean: Thank you for agreeing to be on my committee and for your guidance. Thank you for believing in me and my work. Your insights and encouragement have helped me significantly to improve my work.

To Faculty and Staff: I also want to thank the faculty and staff at the University of Waterloo for creating such a great environment for research and learning.

To Mom, Dad: I am proud to have you as my parents. Thank you for being there for me whenever I needed you. I could not have done this dissertation without your endless support, love and generosity.

To my friends: I want to thank you all for your support and for the good memories we have shared. Your friendship and support have helped me a lot in both good times and bad times.

To Anya: I want to dedicate this dissertation to you. Thank you for being the best and the most supportive daughter anyone can have. Your beautiful smile and hugs always brighten my day. Our little adventures have made my Ph.D. journey so much fun and. Thank you for sacrificing countless hours so that I could work. I could not have done this without your support and love.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

BYM	Besag–York–Mollié Model
BMA	Bayesian Model Averaging
DIC	Deviance Information Criterion
GAM	Generalised Additive Models
GeoMEDD	Geographic Monitoring for Early Disease Detection
GIS	Geographic Information System
GLM	Generalized Linear Model
GLMM	Generalized Linear Mixed Model
GWR	Geographically Weighted Regression
HPM	Highest Probability model
IWLS	Iterative Weighted Least Square
LSE	Land Surface Emissivity
LST	Land Surface Temperature
MAPE	Mean Absolute Percentage Error
MCMC	Markov chain Monte Carlo
MGWR	Multi-Scale Geographically Weighted Regression
RMSE	Root Mean Squared Error
RW	Random Walk
SARS	Severe Acute Respiratory Syndrome
SaTScan	Software for the spatial, temporal and space-time Scan statistics
SEM	Spatial Error Model
SLM	Spatial Lag Model
SVTT	Spatial Variations in Temporal Trends

Chapter 1 : Introduction

1.1. Problem and context

Infectious diseases can often lead to the destruction of lives and impose greater challenges on the global and local public health systems. With the rise of globalization in the modern world and the interconnected population, the recent emergence of infectious diseases involves complex interactions (Threats et al., 2006). The world is becoming more urban, and as a result, the public health community faces numerous challenges in facing emerging infectious diseases. In particular, coordinated global and local approaches are required in densely populated urban areas to prevent and control complex infectious disease issues (Balcan et al., 2010). Big cities can be incubators for new epidemics and can spread in a more rapid manner and become great threats. Rapid detection of emerging infectious diseases and their dynamics is key to minimizing the illnesses, deaths and loss of economy and decreasing the burden of the disease (Neiderud, 2015). In addition to the physical and psychological health challenges in the public health systems, the annual global financial burden attributable to a pandemic can cost over 16 trillion US dollars (Cutler & Summers, 2020). The human and economic costs highlight the need for advanced approaches in infectious disease research (Armitage, 2021).

Coronavirus-19 (COVID-19) is a highly transmissible infectious disease that was first identified in December 2019 in Wuhan, China. COVID-19 is caused by a novel coronavirus, SARS-CoV-2, that can cause acute respiratory illnesses in humans (Bulut & Kato, 2020). Due to the large-scale worldwide outbreak, COVID-19 was declared a global pandemic by World Health Organization (WHO) on March 11, 2020. As of August 15, 2020, COVID-19 contributed

to over 5.85 billion confirmed cases and over 6.4 million deaths worldwide (WHO, 2021). In Canada, COVID-19 was the third leading cause of death in 2020, accounting for 5.3% of all deaths that year. The COVID-19 pandemic contributed to both direct and indirect impacts on global health and the economy across the world. In addition to the many loss of lives, countries worldwide have experienced unprecedented impacts on the labour market (Béland et al., 2020; Demena et al., 2022), economy (Government of Canada, 2021b), supply and demand shocks (del Rio-Chanona et al., 2020), entertainment and tourism industry (Fotiadis et al., 2021) and healthcare.

Infectious disease epidemiology is the study of why and how infectious diseases emerge and spread among different populations and what strategies can prevent or contain the spread of the disease at the population level (Payne, 2017). These infections can spread among humans by contact with members of the households or community or by random contacts from the population (Balcan et al., 2010). Spatial and spatiotemporal analysis of infectious disease incidence is commonly used to understand the spreading patterns and trends and explore the relationship between disease spread and contributing factors (Dongmei Chen et al., 2014). Within communities in an urban area, infectious diseases have been known to be associated with social-ecological processes of neighbourhood and urbanization (Neiderud, 2015). The key step in infectious disease modelling and control is identifying the high-risk areas, spatial transmission patterns and environmental risk factors. Timely detection of outbreaks is important to detect the sources of origin and the direction of the spread patterns, as well as provide important clues to various risk factors (Chowell & Rothenberg, 2018a; Steele et al., 2020).

Alongside the theoretical developments, there have been methodological developments for analyzing spatial and spatiotemporal data of infectious diseases. Earlier studies have used visual maps of cases to identify high-risk areas and associate data (Kamel Boulos & Geraghty, 2020). More recently, spatial and spatiotemporal methods has been used to identify high-risk areas and employ quantitative methods to characterize disease patterns and estimate the relationship between infectious disease and the environment (Kirby et al., 2017). Of particular interest for theoretical and policy developments are the approaches to analyzing the local variations in spatial associations between disease outcomes and sociodemographic and/or environmental risk factors, as well as understanding trends and the spread mechanism of disease outbreak patterns (Caprarelli & Fletcher, 2014; Clarke et al., 1996; Graham et al., 2004; C. Qi et al., 2020; Robertson et al., 2010).

Motivated by non-uniform spatial and temporal patterns of highly infectious diseases, events and research in the past have widely explored the relationships between transmissible disease and the environment (D. Chen, 2014; Franch-Pardo et al., 2020; Nazia et al., 2022; Robertson & Nelson, 2014). Historically, infectious disease epidemiology dates back to 1866, when British physician John Snow was the first to identify that water was the mode of transmission for the deadly disease of Cholera (Snow, 1856). Dr. John Snow observed the spatial variations in cholera cases by plotting the deaths of Cholera on a map and identified the water pump on Broad Street as the source of the cholera outbreak. Since then, various studies have focused on the non-heterogeneous spatiotemporal patterns of infectious disease and clustering to facilitate public health intervention programs (Caprarelli & Fletcher, 2014; Robertson & Nelson, 2014; Saffary et al., 2020; Scarpone et al., 2020; Zhou et al., 2020).

Public health planning for infectious disease outbreaks in urban areas traditionally has focused on identifying areas with high spatial risk (Dummer, 2008). Relationships between the government of urban areas and the control of infectious diseases are not a new concept. Attempts to deal with the pandemic go back to the bubonic plague in the 14th century in Europe and other major epidemics, smallpox, typhoid, yellow fever and Cholera (Cox & Bia, 2017). A study by S Harris Ali and Roge Keil investigated how processes of globalization have affected the transmission and responses to SARS in 2003 within the context of the global cities network and argued that urban governments must be prepared to deal with infectious diseases (Keil & Ali, 2007). One important lesson from the previous SARS epidemic is that we need an efficient global health network to cope with such unanticipated outbreaks since the security of even developed countries depends on the capacity to identify and address potential health threats locally (Banos & Lacasa, 2007).

The neighbourhood effects theories can be used in urban public health policies and programs as they are responsible for working to control disease (Frieden, 2010). It can provide the framework to help shape neighbourhood structure and socio-economic inequalities within the urban environment and see the contexts of how policymakers should approach policy design for current control and in future emerging diseases (Rydin et al., 2012). Limited availability of appropriate diagnostic tests and vaccines can interfere with the large population dense for public health efforts to prevent and control outbreaks (Binder et al., 1999). By observing neighbourhood effects of COVID-19, such as socioeconomic inequalities and environmental processes, planning research can inform the interpretation of disease epidemiology (Quinn & Kumar, 2014). By focusing on population structure, such as ethnic heterogeneity, concentration of immigrants, or

lower education level within a small area, this dissertation assumes that higher concentrations of these variables in neighbourhoods have a positive relationship with COVID-19 transmissions. Previous research shows that an influx of immigrant population can lead to overcrowding neighbourhoods and unsafe housing conditions that are potential health hazards and must be taken into account for safe city planning (Harwood & Myers, 2002; Preston et al., 2009). Focusing on the social-spatial inequality and urban change processes, research has also found that north American cities are experiencing higher levels of income inequality, increasing levels of segregation among low and high-income neighbourhoods (W.-H. Chen et al., 2012).

The development and application of spatial and spatiotemporal epidemiological theories of an infectious disease can help inform, develop, and translate the epidemiological theories applied to public health policy and planning. Infectious diseases persist and thrive depending on how easily they are transmitted and how easily they evolve over time. In urban public health planning, it is important to inquire about what cities must do to prepare for a major health disaster event (Matthew & McDonald, 2006). Cities are ideal targets during the pandemic as urbanization amplifies the infectious disease threats (Madhav et al., 2017). When an outbreak hits a largely populated urban area, it may not be feasible to initiate an intervention in all the areas simultaneously due to resource limitations (Alirol et al., 2011). Past studies have reviewed the important policy implications for identifying hotspots for infectious diseases used in health-related research from the lenses of health promotion and public health (Elliot et al., 2000; Fletcher-Lartey & Caprarelli, 2016; Shrestha & Stopka, 2022). A systematic review published in 2015 listed 80 peer-reviewed articles that studied infectious diseases using spatial methods ranging from respiratory infections (such as SARS and influenza) to intestinal infections (e.g.,

cholera and salmonellosis) to sexually transmitted infections (Smith et al., 2015). The review highlighted the importance of the spatial context tied to infectious disease outbreaks in public health planning, including connections to the origins of key sources, and identifying risk areas and risk factors.

Community-level preparedness and timely and efficient intervention strategies are essential to prevent and control a pandemic (Madhav et al., 2017). For initiatives for COVID-19, it is important to ask if the high-risk areas have adequate surveillance and capacity, whether appropriate medication or vaccines are available, and plan effective quarantine and communication strategies (Sahu & Kumar, 2020). Adequate city planning and surveillance are powerful in improving global health and decreasing infectious diseases (Neiderud, 2015). The capacity of cities to prevent and manage infectious disease spread requires careful consideration. Delays in detecting and responding in high-priority risk areas have led to overburdening local health in many areas worldwide, while effective and timely strategies have helped lower cases in many areas (Khanna et al., 2020). Large urban areas' medical and healthcare systems are often overloaded, and efficient planning to prepare, plan and educate the population is important in a major crisis, and these actions could significantly reduce COVID-19 risks (Neiderud, 2015).

1.2. Rationale

Most recently, with the recent global COVID-19 pandemic affecting both developing and developed countries around the world (Song & Zhou, 2020), the need for theories relevant to COVID-19 at the time of the global public health crisis has heightened to inform appropriately timed responses. COVID-19 is a relatively new disease that has brought unprecedented

challenges around the world. Toronto, the most densely populated city in Canada, has been severely impacted by COVID-19 and has become an epicentre of COVID-19 outbreaks (City of Toronto, 2022a; Detsky & Bogoch, 2020). The local government of Toronto has adopted several control strategies that include emergency lockdowns, stay-at-home orders, increased testing, contact tracing capacities, and closure of in-person schools and non-essential businesses. However, Toronto continued to experience a rise in cases, creating an enormous challenge for public health as well as causing economic and social burdens throughout the pandemic. At least 89% population over the age of 5 years in Canada have received at least one dose (March 28, 2022) (Canada, 2021a). Despite these made by the local and provincial public health official interventions and mitigation measures, Toronto continued to experience substantial COVID-19 incidence and hospitalization rates (CBC News, 2021; CTV News, 2021). As momentum grows to end this global pandemic, understanding the disease trends, detecting hotspots, and identifying important risk factors at the community level is an imperative research effort.

This dissertation was designed to contribute to the ongoing global research effort on COVID-19 by conducting three analytical studies over the course of the pandemic to contribute to public health sciences and to provide targeted solutions. Aligned was the need for understanding the spread dynamics of infectious disease outbreaks in space and time for current or future pandemic control. These areas are important for obtaining critical information to inform policymaking and enhance public health strategies.

While various socio-environmental factors have been known to impact the spread of highly transmissible diseases such as COVID-19, this dissertation seeks to investigate the role of these factors in urban neighbourhoods using Toronto as a case study and explores how various

tools can be used to combat the disease in a dense urban area such as Toronto (Lombardo & Buckeridge, 2006). From past studies, it is clear that planning-led disease prevention initiatives focused on the spatial spread, trends, and socioenvironmental factors can influence the mechanisms that influence infectious diseases as hypothesized by epidemiological disease theories. Therefore, this dissertation plans to explore different spatial methods to investigate trends of risks and outbreaks of COVID-19 in Toronto, the role of social, economic or environmental influences on these outbreak patterns, and to assist with the implementation of place-based pandemic prevention policies and programs (Czabanowska & Kuhlmann, 2021). This dissertation also highlights the importance of neighbourhood-level contexts while analyzing the local spatial and spatiotemporal parts. Additionally, motivated by the recent increase in COVID-19 incidence in winter seasons, this dissertation scrutinizes the effect of temperature and seasonality on COVID-19 incidence in Toronto.

1.3. Literature Review

1.3.1. Urban socioeconomic inequalities

The neighbourhood effect is defined as the processes by which living in a particular neighbourhood influence the well-being of its residents, either collectively or individually (Roosa & White, 2014). A growing body of research has focused on the neighbourhood effects on infectious disease, primarily the impact of neighbourhood socioeconomic status, quality of residence, race, immigrants ethnicity or education on infectious disease outcomes (Bermudi et al., 2021; Bie et al., 2021; Y. Lin et al., 2020; Madhav et al., 2017; Niedzwiedz et al., 2020; Schneider & Machado, 2018). A systematic review of 259 research articles on research on

neighbourhood effects on health in the United States between 1995-2014 identified socioeconomic indicators were the most commonly analyzed neighbourhood variables (Arcaya et al., 2016). The high inequities in economic, social and living conditions in a neighbourhood are often mirrored by inequalities in health and diseases (Sharifi & Khavarian-Garmsir, 2020, 2020). Interest in the neighbourhood effects on health has been motivated by epidemiological studies that seek to explain the patterns of infectious diseases across geographic areas and populations by recognizing that an individual's health is influenced by not only individual characteristics but also by the neighbourhoods to which the individual lives (Arcaya et al., 2016). Examining neighbourhood effects on health outcomes has a practical importance as the public health community looks to place-based intervention plans in order to promote health and health equity in a population (Kivimäki et al., 2021).

Overall, information about infectious diseases and how they spread in the community can help individuals, but information about the impact of neighbourhood effects on COVID-19 outbreaks is crucial for local policymakers and public health authorities. The inequality issues in cities often make it difficult to social distance, thereby undermining some restrictions such as lockdown orders to contain the spread of the virus (Sharifi & Khavarian-Garmsir, 2020). Similarly, in low economic conditions, the need for social interactions for livelihood reasons can also make stay-at-home orders challenging, putting the entire city at risk (Sharifi & Khavarian-Garmsir, 2020). Past studies have found that time-targeted interventions by identifying socioeconomic inequalities may potentially have a profound effect on the population's disease risk (Burton et al., 2010).

Historically, pandemics hit minorities and people in the lower-socioeconomic groups disproportionately (Schneider & Machado, 2018; Wade, 2020). Big cities are often the center of economic opportunities, which can lead to overcrowding of population density, and consequently, neighbourhoods having unsuitably crowded housing can rapidly transmit COVID-19 infections (Ingen et al., 2021). Population from lower socioeconomic groups often suffer from pre-existing health conditions due to economic conditions, limited access to resources or lack of education, increasing their vulnerability to infectious disease outbreaks (Wade, 2020). For example, a recent study by Lo et al. (2021) of US and England adults shows that the community-level risk of positive COVID-19 tests was much higher in ethnic minorities compared to the white population (Lo et al., 2021). Most recently, COVID-19 pandemic studies suggest that the incidences and death rates from COVID-19 showed some disparities among different neighbourhoods, and demographic and socioeconomic factors often explained these variations (Chaudhuri et al., 2021; Wachtler et al., 2020). Some of the key factors to have an influence on the incidence or mortality rates of COVID-19 include income (Abedi et al., 2020; Chaudhry et al., 2020; Cordes & Castro, 2020; Maiti et al., 2020; Sannigrahi, Pilla, Basu, Basu, et al., 2020a), poverty rates (Y. Chen & Jiao, 2020a; Fielding-Miller et al., 2020; Goutte et al., 2020; Richmond et al., 2020; Sannigrahi, Pilla, Basu, Basu, et al., 2020a), education (Abedi et al., 2020; Cordes & Castro, 2020; Goutte et al., 2020; Y. Wu et al., 2020) ethnicity or minority status (Andersen et al., 2021a; Y. Chen & Jiao, 2020a; Cordes & Castro, 2020; Kathe & Wani, 2020; Maiti et al., 2020; Niedzwiedz et al., 2020; F. Sun et al., 2020a; Y. Sun et al., 2020a), immigrant population (Borjas, 2020; Fielding-Miller et al., 2020), and unemployment rate (Goutte et al., 2020; Y. Sun et al., 2020a).

While socioeconomic inequalities have been found to be important contributors to COVID-19 transmissions, substantial contextual variations exist in disease transmissions across different spaces. COVID-19 is still a relatively new disease, and mixed evidence has been reported on the association between COVID-19 and neighbourhood poverty level, housing conditions, unemployment rates, and minority or immigrant status (Booth et al., 2021; Grekousis et al., 2022; Nazia et al., 2022; Wachtler et al., 2020). Different patterns can also be seen across different geographic settings, and factors that are found to be significant in one infectious disease in one urban setting may not be significant in another (Neiderud, 2015).

The city of Toronto has a highly diverse population of 2.99 million (2020) with varying socioeconomic statues across its 140 neighbourhoods (City of Toronto, 2021a). Previous studies have emphasized that people with low socioeconomic or marginalized status might be forced to leave their homes to maintain income or live in congregate settings, which can place them and the neighbourhoods of their residence at a much higher risk during the pandemic (Cordes & Castro, 2020; Y. Sun et al., 2020a; Vaz, 2021). A previous work by Vaz in 2021 (Vaz, 2021) that uses spatial regression models using datasets of the first four months of the pandemic found the socioeconomic vulnerability and population density to be associated with the increasing spread and incidence of COVID-19 in Toronto. This study uses data during the reference periods of 2008, 2011 and 2014 and only observed cases during the first four months when widespread testing options were also unavailable. Another study of Toronto neighbourhoods by Choi et al. in 2021 found that several demographic and socioeconomic factors, such as higher education rates, and lower rates of immigrants (foreign-born residents), were significantly associated with decreasing the number of COVID-19 infections (Choi et al., 2021). This study, however, does

not account for the spatial relationships to help explain the factors behind observed spatial patterns.

1.3.2. Urban land surface temperature and disease transmission

In recent years, changes in urban land use have become more common in urban cities and, therefore, often responsible for the variations in urban land surface temperatures (LST) within different neighbourhoods (Ejiagha et al., 2020). The urban-induced adverse environmental impacts can cause increasing surface temperature (J. Du et al., 2019; Y. Wang et al., 2016; Y. Zhang & Sun, 2019). A combination of factors such as the removal of vegetation, soil moisture, infrastructure development, and concentration of population has been reported in the literature as the drivers of the spatiotemporal changes in LST (Senanayake et al., 2013).

Past studies have found day-to-day LST variations within different neighbourhoods within a city (Anderson et al., 2018; Buyantuyev & Wu, 2010; Rinner & Hussain, 2011; Tam et al., 2015). A study conducted in Phoenix, Arizona and Dallas, Texas, in the United States concluded that the urban areas that are susceptible to a high LST are primarily occupied by vulnerable population groups (Moss & Kar, 2020). Furthermore, temperature is often a significant risk factor for infectious diseases because a certain temperature can help the virus evolve rapidly (Tobías & Molina, 2020). Therefore, urban communities with a higher land surface temperature can have a relatively higher risk of infectious disease transmission.

A few studies have explored the impact of LST on COVID-19. A study by Das et al. in Kolkata, a densely populated city in India, found LST to have an influential role in the spatial clustering of COVID-19 in the communities (Das et al., 2020). A study in all USA counties by Johnson et al. found a strong negative influence of nighttime land surface temperature (LST)

with COVID-19 using low-resolution Moderate Resolution Imaging Spectroradiometer (MODIS) images of 2020 (Johnson et al., 2021). A study by Hassan et al. identified a strong positive relationship between COVID-19 and LST in the most populated city in Bangladesh (Hassan et al., 2021). This study showed that a 1°C increase in LST is linked with a 36.1% increase in COVID-19 incidence rates. Even though these studies have attempted to understand the relationship between urban land surface temperature and COVID-19, most of these studies have used short temporal periods and lower-resolution images without any atmospheric corrections. The results were contradictory in different geographic areas, leaving a gap in understanding the impact of surface temperature on COVID-19 transmission within neighbourhoods. To the best of my knowledge, no studies have looked into the impact of urban temperature or LST on COVID-19 transmission in Toronto or in the Canadian context, and the understanding of spatiotemporal variations in urban land surface temperature on COVID-19 transmissions across neighbourhoods may contribute to the understanding of COVID-19 spatial epidemiology.

1.3.3. Outbreak spread dynamics

During an epidemic, observing the movement of outbreaks across neighbourhoods in space and time can allow identifying possible sources and directions of the outbreaks. A few earlier studies that mapped infectious disease outbreaks mostly involved comparing multiple maps to compare the outbreak shifting visually. For instance, a few studies evaluated the spread of COVID-19 over space and time by visually comparing multiple maps to observe the spatial distribution of new or cumulative cases using multiple time points during the early stages of the pandemic

(Cuadros et al., 2020; Y. Feng et al., 2020; Gao et al., 2021; H. Yu et al., 2021). Cuadros et al. observed the COVID-19 case burden over three temporal dates (March 30, April 20 and May 10) in 2020 and visually compared the spread patterns and dynamics in Ohio, US (Cuadros et al., 2020). Feng et al. visually explored the spatiotemporal spread pattern cases of COVID-19 in China by showing the new COVID-19 cases reported each week from January 17, 2020, to March 20, 2020 (8 temporal dates) in and around Wuhan (Y. Feng et al., 2020). A study by Gao et al. observed the spatiotemporal distribution in Hubei province, China, by visually comparing COVID-19 cases on four separate dates during the earliest stage of the pandemic. Similarly, another study by Gianquintieri et al. in Lombardy, Italy, used choropleth mapping to display the beginning of COVID-19 onset from emergency calls while not including details about the spread mechanism (Gianquintieri et al., 2020). Spread patterns were also modelled in other infectious diseases, such as the 2009 pandemic influenza outbreaks. A transmission model by Gog et al. showed the spatial spread of the 2009 pandemic influenza but lacked further details about the starting point and the direction of movement of the outbreaks (Gog et al., 2014). However, these models lack visualizing disease outbreaks by including temporal structure, dynamics and direction of the outbreak in a single map. Additionally, earlier studies have not assessed the sources of infection datasets in dispersed areas.

1.3.4. Spatial and spatiotemporal heterogeneity

The hotspot is an emerging and dynamic concept that can change based on the spread of the disease and help identify significant risk regions, especially during the COVID-19 pandemic (Islam et al., 2021). Significant knowledge of spatial clusters can provide a baseline approach for health researchers to understand disease occurrence and generate hypotheses about the

etiological factors of disease (Kirby et al., 2017). A few studies have emphasized the importance of identifying locations of hotspots which can help make emergency measures using the available resources (Desmet & Wacziarg, 2020; S. S. Lee & Wong, 2011, 2011). Some studies have analyzed the spatial dynamics of the COVID-19 epidemic that is experienced on a global scale (Franch-Pardo et al., 2021; Z. Huang, 2021; Shariati, Mesgari, et al., 2020).

Facilitated by Geographic information systems, various studies in the literature found uneven spatiotemporal patterns between COVID-19 and urban neighbourhoods (X. Chen & Quan, 2021; Nazia et al., 2022; Paul et al., 2020; Verma et al., 2021). For instance, Kang et al. observed the spatial spread of COVID-19 using data from January 16 to February 6, 2020, at the regional scale in China that found the presence of spatial dispersion using Moran's *I* statistic (Kang et al., 2020). Another study by Ghosh and Cartone identified strong spatial clusters in Italy between February 24 to June 26, 2020 (Ghosh & Cartone, 2020). A systematic review by Nazia et al. has identified that the majority of the earlier work in the area of COVID-19 spatial epidemiology primarily focused on identifying hotspots of COVID-19 in China and the USA and lacks spatial studies that identified spatiotemporal variations in the Canadian context (Nazia et al., 2022). This review has also identified that most of the hotspots are conducted at a global or regional scale, lacking community-level analysis, particularly in Canada. An earlier work using Toronto neighbourhood datasets by Vaz in 2021 used local spatial autocorrelation methods to find that COVID-19 cases were not uniformly distributed in Toronto neighbourhoods during the first four months of the pandemic (Vaz, 2021). However, this study only focuses on the earliest part of the pandemic and uses a spatial approach without taking temporal factors into account.

1.4 Dissertation objectives and gaps to fill

This dissertation has three primary objectives. The first objective aims to advance the understanding of the disease epidemiology of COVID-19 in Toronto. In particular, the dissertation investigates the COVID-19 spread patterns and the spatial correlations between urban characteristics and COVID-19, taking into account the urban socio-environmental aspects. The second objective is analytical and uses applied and novel spatial and spatiotemporal models to strengthen inference of small-area level COVID-19 and risk factors, identify high-risk areas and trends and visualize outbreak spread dynamics. The third objective is public health policy-oriented which focuses on providing quantitative information for policies for preventing infectious diseases with consideration of the spatial relationship between urban characteristics, infectious disease and the management of public health. Table 1.1 describes the specific research objectives for each of the three manuscripts of this dissertation. The details are further discussed in the dissertation below.

Table 1.1. Specific research questions and objectives for each manuscript.

Chapter #	Objective	Research Questions
Chapter 3	Analyze the spatial and spatiotemporal clusters of COVID-19 and identify globally and locally variable socioeconomic drivers of the COVID-19 incidences.	1. Where are the high-risk areas located? 2. What are the significant global socioeconomic risk factors of COVID-19? 3. What are the local neighbourhood-level variations in the association between COVID-19 and socioeconomic risk factors?
Chapter 4	Analyze spatiotemporal variations of COVID-19 and identify temporal,	1. Where are the high-risk areas located? 2. What are the important socioeconomic risk factors of COVID-19?

	spatiotemporal trends and identify local socioeconomic and environmental (LST) risk factors.	3. Is there a relationship between COVID-19 and Land surface temperature (LST) while adjusting for socioeconomic and demographic factors? 4. What are the temporal trends of COVID-19 transmission? 5. What is the spatiotemporal trend of COVID-19 transmission (e.g., stable, increasing or decreasing)?
Chapter 5	Visualizing the spatiotemporal spread dynamics of COVID-19 outbreaks and prioritizing risk areas.	1. What are the spatiotemporal spread dynamics of COVID-19 outbreaks? 2. Where are the high, medium and low-priority hotspots located?

1.4.1. Epidemiological Objectives

This dissertation asks and answers several questions that advance the understanding of the epidemiology of COVID-19 and explains the dynamics and relationship between COVID-19 and the urban environment. The findings advise the design of the geographic models and recommend public health prevention and intervention plans for COVID-19 (Figure 1.1). Identifying disease clusters and understanding the driving factors for these clusters using spatial analytical approaches can provide us with a more realistic view of the issue in Toronto compared to the traditional simple maps (Cordes & Castro, 2020). Even though Toronto remains at heightened risk in this pandemic, spatial studies on the burden of COVID-19 in Toronto have so far been very limited. Understanding the impacts of neighbourhood effects on COVID-19 outcomes is essential to further advance the understanding of COVID-19 epidemiology in Toronto and help policymakers design more appropriate protection and responses.

Similar to other infectious diseases, COVID-19 transmission is more frequently observed in urban areas. Across the three manuscripts, I asked a common question about where the high-risk areas are located in the study area using multiple spatial modelling techniques. It contributes to current knowledge on the geography of infectious diseases of how COVID-19 risks can be heterogenous in space and space-time across urban neighbourhoods and identifies COVID-19 risk variations in Toronto.

Each of the three manuscripts has its own objectives that advance the current understanding of the epidemiological theories of COVID-19 in Toronto (Figure 1.2). Prior studies in infectious disease research, including COVID-19, have found urban socioeconomic inequalities to significantly influence COVID-19 (Alirol et al., 2011; Nazia et al., 2022; Wachtler et al., 2020). However, the results were diverse and often contradictory in different geographic areas, leaving a gap in understanding the impact of neighbourhood-scale socioeconomic inequalities on COVID-19 transmission in Toronto. In Chapter 3 and Chapter 4, I study how socioeconomic inequalities across the neighbourhoods globally influence COVID-19 in Toronto. In Chapter 3, I further explore the local variations in the relationships between COVID-19 and socioeconomic vulnerabilities across different neighbourhoods. Compared to the prior study in Toronto by Vaz (2021), this chapter explores using a longer temporal period as well as uses a more recent census dataset which may improve the reliability of the association between disease risk and the socioeconomic variables (Vaz, 2021).

While prior studies on COVID-19 have primarily focused on identifying high-risk areas, to the best of my knowledge, no study has yet identified whether spatiotemporal COVID-19 patterns are stable or show an increasing or decreasing trend over time in a small area, including

in Toronto. This dissertation fills this gap by modelling the temporal or spatiotemporal trends of COVID-19 in Toronto (Merow & Urban, 2020). In Chapter 4, I specifically asked what were the temporal or seasonal trends and to what degree the spatiotemporal trends varied for each neighbourhood. Even though in an urban environment, the land surface temperature is known to vary (S. Wang et al., 2018), and temperature is often known to impact the transmission mechanism of infectious diseases (Chua et al., 2022; Rohr & Cohen, 2020), only a few studies considered the impact of urban neighbourhood variations in temperatures on the surge of COVID-19 outbreaks (Mukherjee and Debnath 2020; Aboubakri et al. 2022; Moss and Kar 2020). To the best of my knowledge, no study has looked at the spatiotemporal impacts of LST in Toronto. In Chapter 4, I also asked how the land surface temperature in Toronto influences the spatiotemporal patterns of COVID-19. Identifying land surface temperature as a risk factor provides evidence of the impact of the spatiotemporal variations in neighbourhood-level LST temperature on COVID-19 risk. This work accounts for the theories of infectious diseases in a Canadian urban city and extends past work of LST that has used fewer temporal points by using weekly data on COVID-19 and land surface data over 79 weeks.

Literature and method reviews (Sections 1.3.3 and 2.3) identified that previous works in identifying the outbreak spatiotemporal epidemiology primarily compared maps of new COVID-19 cases from multiple time points in a spatial area to compare the outbreak progression visually (Cuadros et al., 2020; Y. Feng et al., 2020; Gao et al., 2021; H. Yu et al., 2021). No studies at the intra-urban scale have modelled the spatial patterns of dispersions of COVID-19 within neighbourhoods or identified the possible source, direction and magnitude of these dispersions within neighbourhoods. This dissertation fills this gap by analyzing the spatiotemporal modelling

of the contagion inside cities on the neighbourhood scale, where interpersonal contact occurs more intensely, creating complex relationships between the population and the environment (Perles et al., 2021). In chapter 5, I asked how the spatiotemporal outbreak dynamics patterns varied across the neighbourhoods, particularly whether a large part of the first outbreaks was reproduced in close communities within two weeks. This work advances urban neighbourhood scale research by displaying the spatiotemporal distributions of the contagions inside the cities within the temporal incubation period and identifying an index or source neighbourhood that can help effectively respond to outbreaks at the source and provide the ability to understand how to respond to them. Additionally, this work advances the knowledge of the sources of infection in dispersed neighbourhoods.

1.4.2. Analytical objectives

This dissertation uses one novel and several applied analytical methods to explore the complex relationships between COVID-19 spatial epidemiology and the urban environment. Each of the three manuscripts has its own methodological framework to each address the specific research objectives of this dissertation (Figure 1.1). Chapter 3 employs a local and a global spatial regression model to identify the significant socioeconomic risk factors to assess the overall risk for Toronto. In Chapter 4, in order to identify the spatiotemporal and temporal trends, Bayesian spatiotemporal models with a set of observed covariates and unobserved spatial random effects, temporal random effects and three different types of space-time random effects are explored. The weekly LST data (79 weeks) at the neighbourhood scale is corrected using an atmospheric correction method and fitted into the model. Chapter 5 applies a novel model for

visualizing the spatiotemporal dynamics of the outbreaks in a single map by identifying the source and dispersion neighbourhoods, magnitude and direction of the outbreaks.

This dissertation applied several spatial models to identify high-risk areas in Toronto. Each chapter uses different methods, such as popular test-based methods SaTScan in Chapter 3, Bayesian spatiotemporal models to identify the area-specific relative risk in Chapter 4 and a model that uses disease persistence and incidence rate to identify hotspots in Chapter 5. Hotspot prioritization using outbreak duration or disease persistence has been previously used for other infectious diseases (Debes et al., 2021; Ngwa et al., 2021) but has not been explored in the case of COVID-19. Chapter 2 provides a more detailed discussion of the methodologies used in this thesis. The findings from these hotspot methods are further compared and discussed in Chapter 6.

1.4.3. Public Health Planning objectives

The understanding of COVID-19 disease epidemiology in an urban environment using spatial tools can provide information for the planning, implementation and evaluation of existing COVID-19 prevention plans and policies. This dissertation provides information in each of the three manuscripts that could be utilized for COVID-19 prevention, mitigation and intervention planning and policies. The public health planning contributions of each manuscript also inform theories and hypotheses focused on the spatial epidemiology of COVID-19.

Chapters 3 and 4 explore the relationship between urban neighbourhood socioeconomic vulnerabilities and the impact on COVID-19 transmissions. These socioeconomic determinants of COVID-19 may provide important clues about the spatial drivers of COVID-19 to identify the localities and allow the policymakers to establish disease surveillance based on the significant

socioeconomic drivers that were influencing disease risk in individual neighbourhoods. The research findings may help develop plans for decreasing the socioeconomic inequalities in high-risk neighbourhoods to mitigate the disease risk. In addition, the findings may help identify areas where various health programs, such as COVID-19 awareness and education programs, improve public knowledge and awareness of the disease.

In Chapter 4, I explore whether high LST in a neighbourhood favours the coronavirus transmission rates. The findings can provide information to the policymakers to identify urban surface temperature as a significant influencer and incorporate it into COVID-19 predictions in the neighbourhoods. Chapter 4 also identifies the temporal and seasonal trends of COVID-19 that can allow public health and policies to prepare in advance for future outbreaks. Additionally, the spatiotemporal trends in the neighbourhood may allow policymakers to further evaluate existing strategies to focus on the neighbourhoods that show consistently increasing trends of disease risk.

In Chapter 5, I demonstrate the spatiotemporal spread patterns within neighbourhoods, which can be used to identify the possible sources of infection to prevent further spread. This method can be adopted by public health officials to quickly demonstrate the spread patterns during a pandemic. Close attention to the index and its nearby neighbourhoods could have controlled the spread of the disease in our study area. The visualization may be useful for public health planners to evaluate the impact of existing intervention plans, showing where intervention efforts are successful and where they are not and, therefore, could help identify the barriers to successful intervention programs.

A public health program for controlling an outbreak requires targeted intervention, as there is always a logistical concern for an untargeted intervention. The targeted intervention requires identifying the high-risk areas where special attention is needed. In this dissertation, I explore several spatial models (Chapters 3-5) to identify the high-risk areas for COVID-19 in Toronto, where special spatial attention is needed for controlling the outbreaks, and the findings and applications of the methods in public health sciences are discussed in Chapter 6.

1.5. Dissertation structure

This dissertation explores the spatial and spatiotemporal dynamics of disease risks and outbreaks and the connections between COVID-19 and the socioeconomic and environmental factors in an urban environment, using Toronto as a case study. The overall goal is to help understand the spread dynamics of COVID-19 and the influencing factors of the spread, which can provide evidence of strategies to control COVID-19 in Toronto and provide methodological frameworks for future epidemic control. This dissertation is composed of six chapters. The introductory chapter (Chapter 1) outlines the problem context for the work, establishes its purpose and objectives, provides a review of the literature, and explains how the manuscripts presented in the body address the purpose and objectives. Chapter 2 introduces and explains the methodologies used across the three research studies to address the objectives of this dissertation.

The following three chapters (Chapters 3 to 5) each feature three manuscripts prepared for peer-reviewed journals that have been either published or submitted for publication on the subject of spatial and spatiotemporal epidemiology of COVID-19. Chapter 3 analyzes the spatiotemporal clusters and the socioeconomic drivers of COVID-19 in Toronto neighbourhoods. The focus was to explore the neighbourhood trends (i.e. how COVID-19 risk varied over the neighbourhoods) and identify the local and global socioeconomic drivers that influence these

trends. In addition to similar past studies, this study further explores the neighbourhood-level variations in socioeconomic risk factors. This study sheds light on how socioeconomic differences can amplify the COVID-19 risk in an urban area and provides policy implications for improving intervention strategies.

Chapter 4 uses a Bayesian hierarchical spatiotemporal modelling approach to explore the neighbourhood-specific COVID-19 risks, spatiotemporal and temporal risk trends and identify important risk factors in an urban environment. This modelling approach quantifies the uncertainties to predict the temporal and spatiotemporal risks of COVID-19. The study advances the understanding of spatiotemporal variations in urban land surface temperatures and socioeconomic conditions influencing COVID-19 risks and identifies high-risk areas where interventions can be targeted.

Chapter 5 applies a novel method to visualize the COVID-19 spread dynamics in an urban area and prioritizes high-risk areas for intervention. Our study offers a new approach to advance the understanding of the spatiotemporal dynamics of infectious disease outbreaks, which would facilitate the intervention program for controlling the spread of the outbreaks.

Chapter 6 provides a discussion and concludes this dissertation by summarizing the major findings, comparing and commenting on the findings across the three studies, and highlighting major contributions. The limitations of this dissertation and future research directions are also discussed in this chapter.

Chapter 2 Methodology

2.1. Spatial epidemiology

The current research is from the domain of spatial epidemiology, which can be defined as the description and analysis of spatial and temporal variations in disease with the target formulation of a hypothesis about the etiological factors of that disease (Elliott & Wartenberg, 2004). Generally, the field of spatial epidemiology can be divided into three principal areas: disease clustering, spatial regression or geographic correlation and disease mapping methods (Elliot et al., 2000). The disease clustering methods identify the high-risk areas for a particular disease in a region, while the spatial regression methods analyze the relationship between disease risk and influencing factors on a particular geographic scale. Finally, the disease mapping methods produce descriptive maps that can provide a visual summary of disease occurrence in a region (Elliot et al., 2000).

2.2. A review of spatial methods used for infectious disease modelling

In an urban setting, geography, neighbourhood effects and the ecological aspects of infectious diseases describe the processes contributing to outbreaks within areas (Scott & Carrington, 2011). The neighbourhood effects in studying public health and infectious disease outbreaks are the phenomena in that the health outcome in one space is affected by the health outcomes in its surrounding space (Fullilove, 2003). The transmission of infectious diseases from human to

human is a socio-spatial process, and the health outcome is spatial-dependent (C.-H. Lin & Wen, 2022). Taking these spatial dimensions is crucial when considering the methods to observe the epidemiology of infectious diseases (C.-H. Lin & Wen, 2022). In the following parts of this chapter, I provide a brief review of a few geospatial methods that have been used for infectious disease mapping, discuss the main methods adopted in the three manuscripts, and describe the scale, datasets and case study while highlighting the applicability of using Toronto as a case study.

A number of test-based and model-based methods are available to understand the spatial heterogeneity, dynamics and risk factors of infectious diseases. Geospatial methods of spatial autocorrelation, hotspot and clustering analysis using GIS can be an effective tool in spatial epidemiology and provide a way to analyze the geographic patterns and etiology of disease. Ecologic regression models are commonly used to assess the association between health outcomes and risk factors at the area level. Spatial relationships in infectious disease outbreaks and surveillance data can be analyzed through various modelling approaches to create an informative representation of environmental processes in geographic space.

Kulldorff's spatial scan statistic method implemented in SaTScan software is the most widely used statistical method for infectious disease high-risk areas or for cluster detection (Unkel et al., 2012). In SaTScan, the study region is scanned using a number of circular or elliptical windows to identify whether a group of contiguous areas has the most significant excessive risk in the region (Kulldorff, 1999). A number of studies have used the spatial (Andersen et al., 2021b; Benita & Gasca-Sanchez, 2021; Chakraborty, 2021; Deguen & Kihal-Talantikite, 2021; Escobar et al., 2021; Ladoy et al., 2021; Ullah et al., 2021) and space-time scan statistic (C. Qi et al., 2020; Rosillo

et al., 2021; Tyrovolas et al., 2021; Y. Wang et al., 2020; Xu & Beard, 2021; J. Zhang et al., 2021) using SaTScan to identify spatial and spatiotemporal clusters of COVID-19. For example, Hohl et al. identified weekly clusters of COVID-19 in the USA during the earlier stages of the pandemic using the Poisson-based space-time statistic using SaTScan to facilitate disease surveillance and to improve resource allocation (Hohl et al., 2020). Similarly, Andrade et al. used SaTScan to detect space-time clusters of COVID-19 in Sergipe, Brazil (Andrade et al., 2020). The major advantage of this method is that it overcomes the multiple-testing issues common to many spatial analyses and can incorporate the heterogeneous population information and visualize both global and local tests for clustering (Sherman et al., 2014). A study by Alkhamis et al. used Kulldorff's multivariable permutation scan statistic (MPSS) implemented in SatScan that accounted for socioeconomic variables to understand the spatial and temporal patterns of COVID-19 pandemic in Kuwait (Alkhamis et al., 2020). A limitation of this study included the fact that the method does not account for the direct effect of the population in the analysis.

Spatial autocorrelation statistic measures the degree to which spatial features and associated rates or values have a tendency to cluster together in space and thus exhibit similar disease rates or are dispersed (Kirby et al., 2017). Alternatively, the local Moran's *I* identify where the hotspot or cold spot clusters are occurring as well as the shape and significance of these clusters (Anselin, 1995). These global and local Moran's *I* was also widely used by many researchers to identify the presence of spatial heterogeneity and to identify hotspots of COVID-19 using GeoDa software packages (Bilal et al., 2020; Han et al., 2021; H. Li, Li, et al., 2020; Saffary et al., 2020; Xie et al., 2020; Xiong et al., 2020; L. Ye & Hu, 2020). For example, Kang et al. determined that there is a spatial dispersion with a significant global Moran's *I* statistic

suggesting the presence of spatial heterogeneity of COVID-19 between January 16 to February 6, 2020, for 31 regions in China (Kang et al., 2020). In another study, Islam et al. identified where the risk of the COVID-19 epidemic is high with hotspots in Bangladesh using the spatial autocorrelation analysis (Islam et al., 2021).

Spatial regression models can be fitted into GeoDa software to estimate the relationship between the response and explanatory variables (Anselin et al., 2006). These global models include the spatial lag model (SLM) or the spatial error model (SEM), which assumes a spatial dependency between explanatory and response variables (Anselin et al., 2006). These spatial regression models have been extensively used in COVID-19 studies (Feinhandler et al., 2020; Hu et al., 2020; B. Kim et al., 2021; Kulu & Dorey, 2021; Vaz, 2021; You et al., 2020). These global spatial regression models are implicitly based on assumptions that the relationship between dependent and independent variables is stationary over space and the relationship does not vary over space (Anselin et al., 2006).

Geographically weighted regression (GWR) is a local regression method that has often been found to identify the local level association between the outcome variable of interest and potential disease risk factors (Brunsdont et al., 1998). GWR has been extensively used for understanding infectious disease epidemiology (C. H. Lin & Wen, 2011; F. Liu et al., 2020; Mohammadinia et al., 2019; Nazia et al., 2022; X. Wu & Zhang, 2021; Y. Ye & Qiu, 2021; Zheng et al., 2022). The advantage of the GWR method is that this technique will allow us to understand the spatially varying relationship between variables by allowing regression coefficients to differ across space. GWR also has the benefit over global methods as it accounts for spatial autocorrelation in the residuals (Cho et al., 2009). It should be noted that it is possible for a variable

that is not found significant from a global regression method to be found significant at the local level using the GWR method (Cho et al., 2009).

Multi-scale Geographically weighted regression (MGWR) method was recently developed (2017) by Fotheringham et al., which is an extension of the traditional GWR method (Fotheringham et al., 2017). In contrast to the GWR method, the MGWR method allows for studying the relationship between variables at different scales (H. Yu et al., 2020). The MGWR has been applied in understanding infectious disease epidemiology (Y. Chen et al., 2021; Donša et al., 2021; Grillet et al., 2010; C.-H. Lin & Wen, 2022; Mansour et al., 2021; Pranzo et al., 2022; Xie et al., 2021). The strength of the MGWR method is that it obtains a set of optimal covariate-specific bandwidths in which each bandwidth indicates the spatial scale at which a factor impacts the outcome variable (Fotheringham et al., 2017). Mansour et al. used the MGWR method to observe the local effects of socioeconomic determinants on COVID-19 incidence rates in Oman (Mansour et al., 2021). Similarly, Dutta et al. used the MGWR method to observe the role of different factors in the outbreak of COVID-19 in India at the local level.

Generalized Linear Mixed Modeling (GLMM) provides a regression model-based framework to model disease counts or rates using any of the exponential family of statistical distributions (Robertson et al., 2010). GLMM combines fixed effects for explanatory variables with a random effect that accounts for spatial dependence and provides an optimal spatial prediction (Hefley et al., 2017). The drivers of disease emergence are often organized hierarchically through administrative units (Auchincloss et al., 2012). The GLMM approach can easily incorporate hierarchical effects and, therefore, an ecological relationship that is structured hierarchically, impacting disease emergence, such as climatic factors, can be represented and

accounted for (Robertson et al., 2010). The expected values are adjusted towards the mean where fewer people are available in an area, accounting for the small-number problem in spatial epidemiology (Robertson et al., 2010). GLMM models have been used in infectious disease modelling in the past. For example, Kindi et al. used a GLMM model to observe the association between COVID-19 cases and different demographic and socioeconomic factors of COVID-19 in Oman and predicted early incidence and infection rates (Kindi et al., 2021). Another study by Chien et al. used a Poisson-based GLMM model with a spatial function to evaluate the impact of weather variability using meteorological factors such as temperature, relative humidity and precipitation on COVID-19 (Chien et al., 2021).

Bayesian hierarchical models can be used for modelling disease risk and the hierarchical spatial and spatiotemporal relationships between disease cases and the ecological drivers. Analysis of disease emergence in a Bayesian framework has been extensively used in many infectious disease mapping studies (Bermudi et al., 2021; Blangiardo, Cameletti, et al., 2020; Briz-Redón et al., 2022; DiMaggio et al., 2020; Gayawan et al., 2020; Jalilian & Mateu, 2021; Johnson et al., 2021; Lima et al., 2021; Millett et al., 2020; Ngwira et al., 2021; Yang et al., 2021). Bayesian methods are centred around inference on unknown area-specific relative risks based on the observed data and a prior distribution. The earlier works of (Besag et al., 1991), (Clayton & Bernardinelli, 1996), and (Bernardinelli & Montomoli, 1992) have significantly influenced a shift from frequentist to Bayesian approaches to disease mapping (MacNab, 2022).

The basic Bayesian methods can incorporate space and time dependencies. For the spatial autocorrelation term, a conditional autoregressive (CAR) is often suggested for modelling spatial autocorrelation. The Besag-York- Mollié (BYM) model (Besag et al., 1991), a lognormal

Poisson model in the Bayesian framework developed for disease mapping, is more commonly used for the spatial component of Bayesian hierarchical disease modelling (MacNab, 2022). A BYM model includes both an ordinary random-effects component to account for non-spatial heterogeneity and a component for spatial smoothing (Morris et al., 2019). DiMaggio et al. used a Poisson-based Bayesian GLMM model that incorporates a spatial random effect to model the high risk of COVID-19 in New York City while adjusting for various demographic, socioeconomic and racial factors (DiMaggio et al., 2020).

Bayesian spatiotemporal models allow incorporating the spatial and temporal information into analysis in ways that reflect the influences of space and time as well as the interactions of space and time on the phenomenon of interest (Knorr-Held & Besag, 1998; MacNab, 2022). These models are often motivated to facilitate borrowing information and to smooth risk areas within and between space and time (MacNab, 2022). For example, a study by Jalilian and Mateu used a Bayesian Poisson-based spatiotemporal model with spatial, temporal random effects and temporal covariates to analyze relative risk variations of COVID-19 focusing on Spain, Italy and Germany (Jalilian & Mateu, 2021). Another study by Bermudi et al. used a Poisson-based Bayesian hierarchical model that incorporates a spatial structure (BYM model), a temporal structure (Random Walk model) and a space-time interaction term proposed by Knorr-Held (2000) to estimate the relative risk of COVID-19 mortality in the city of Sao Paulo (Bermudi et al., 2021; Knorr-Held, 2000).

Bayesian methods are often attractive as they allow expert and local knowledge of disease processes to be incorporated with the adding of prior distributions. Generally, a sensitivity analysis of the priors for the model parameters is recommended (Haining & Li, 2020;

MacNab, 2022; MacNab & Gustafson, 2007). A further complexity of Bayesian models is that MCMC methods are required for generating posterior distributions that are computationally very demanding (Best et al., 2005). The technical aspects of model fitting require advanced statistical training (Best et al., 2005).

Often disease datasets contain a high number of zero counts of disease cases across space or in space-time and show greater variability in data. A common task in applied statistics is to choose a model to fit these excessive zero counts and overdispersion (Neelon et al., 2013). For count data with an excess of zeros and overdispersion, zero-inflated regression models such as zero-inflated Poisson (C.-S. Li et al., 1999) and negative binomial hierarchical models have been used in the past. An alternative Poisson hurdle model have been developed those deals with the excess zeros and over or under-dispersed data (Corpas-Burgos et al., 2018). Gayawan et al. used a two-component Bayesian geo-additive hurdle Poisson model structured spatial and Spatiotemporal effects to simultaneously analyze the zero counts and the frequency of occurrence of COVID-19 cases in Africa (Gayawan et al., 2020).

The Bayesian Model Averaging (BMA) technique (Raftery et al., 1997) estimates all candidate models and computes a weighted average of the estimates while taking the uncertainties of the models into account. For example, Olmo and Sanso-Navarro applied the Bayesian model averaging (BMA) technique using a Poisson GLM model to model COVID-19 in the city of New York at the zip code level (Olmo & Sanso-Navarro, 2021). BMA is useful in cases when researchers are interested in a particular parameter but uncertain about how the parameter relates to the observations (Hinne et al., 2020). A limitation of BMA is that this method will identify one single model as the true model and may not identify the correct model

if the true data-generating model is not considered as one of the models (Vehtari et al., 2019; Vehtari & Ojanen, 2012).

Geostatistical interpolation or smoothing methods such as Kriging, Thiessen polygon, or Inverse Distance Weighted (IDW) interpolation are also used to improve the accuracy of infectious disease rates for small area analysis (Berke, 2004). These geostatistical interpolation methods are used to predict disease rates at an unmeasured location to create smoothed rates to improve disease rates for small areas with fewer observations (Berke, 2004). For example, Nasiri et al. (Nasiri et al., 2021) used the IDW method to create interpolated maps of infected COVID-19 patients across Tehran and Ramírez, and Li (Ramírez & Lee, 2020) used the IDW algorithm to interpolate and create COVID-19 risk surface map in USA counties at five-time points. Similarly, Arif et al. (Arif & Sengupta, 2021) used the Thiessen polygon smoothing technique to create polygon maps to show the spatial distribution of COVID-19 cases across Southern Indian states. However, strong smoothing or over-smoothing of the disease mapping model can often lead to conservative risk estimates, and therefore a small number of false-positive findings can be common in these interpolation methods (Blangiardo, Boulieri, et al., 2020). It can involve the idea of discontinuous geographic boundaries and hide important information about variation in rates in a small geographic area (Blangiardo, Boulieri, et al., 2020).

Earlier spatial statistical disease mapping primarily focused on the development of empirical Bayes models to estimate the procedures for smoothed maps of disease rates or relative risks (Liao & Brookmeyer, 1995; Marshall, 1991; Tsutakawa et al., 1985). A number of studies have used local empirical Bayesian smoothing technique to get spatially smoothed rates of COVID-19 incidence in each spatial unit of analysis (Andrade et al., 2020; Castro et al., 2021;

M. C. Ferreira, 2020; MacNab, 2022; Ramírez-Aldana et al., 2020; Raymundo et al., 2021; A. P. de S. C. Silva et al., 2021) that estimates the prior distribution from the data (Maritz & Lwin, 2017). The main limitation of the empirical Bayes estimation is that the empirical Bayes standard errors typically underestimate the associated estimation uncertainties (Ainsworth & Dean, 2006; MacNab, 2022; MacNab et al., 2004).

Several new spatial tools or methods have also been recently developed for infectious disease surveillance. For example, Geographic Monitoring for Early Disease Detection (GeoMedd) (A. Curtis et al., 2020) is a recently developed automated clustering method through integration of a spatial database and two types of clustering algorithms. GeoMedd uses incoming test data to provide multiple spatial and temporal clusters by connecting cases with various spatial and temporal thresholds. A few studies have adopted this method to investigate COVID-19 clusters. For example, Curtis et al. (A. J. Curtis et al., 2022) and Miller et al. (A. K. Miller et al., 2022) have used this method to identify clusters of COVID-19 in the Cuyahoga County of Ohio, USA. However, there are some limitations associated with GeoMedd as it requires an individual with a spatial data background to look at the map and understand its dynamism (A. Curtis et al., 2020). It has also been argued that over-automation may have lessened the effectiveness of GeoMeDD (A. Curtis et al., 2020).

Another recently developed spatial statistical method GeoDetector has been used in a few ecological regression studies that quantified the associations between COVID-19 disease risk and the potential risk factors (L. Wang et al., 2021; P. Wang et al., 2021; X. Wu et al., 2020; Xie et al., 2020). The GeoDetector method is not restricted to the assumptions of traditional statistical methods and involves complex processes of parameter settings. Since GeoDetector only works

with the independent variable as a categorical variable, converting continuous data to categorical data may affect the analysis results due to the singleness of the adopted methods (W.-J. Jia et al., 2021).

Rawat and Deb proposed a model structure in a Bayesian framework that includes a separable Gaussian spatial-temporal process model implemented with a random error structure and an additive mean structure and estimates the spatial relative risk of COVID-19 (Rawat & Deb, 2021). This proposed approach would estimate the short-term and long-term predictions for infectious disease for any geographic location, even if it were unobserved in the data. The proposed new Bayesian method would require appropriate approximations to ensure feasible computation.

Nazia et al. in 2022 have further provided a systematic review of the various frequentist and Bayesian spatial and spatiotemporal methods used to model COVID-19 risks and ecological regression (Nazia et al., 2022). This systematic review provides a detailed overview of the common and proposed infectious disease modelling methods that have been used to model the COVID-19 risk during the pandemic.

2.2.1. Infectious vs non-infectious disease spatial modelling

Non-infectious diseases refer to diseases not transmitted from one person to another, while infectious diseases refer to diseases that can be transmitted from one person to another. At present, most of these spatial epidemiological methods are primarily focused on infectious disease surveillance (Lombardo & Buckeridge, 2006). Spatial epidemiology has become an essential tool to fight against infectious epidemics such as malaria (Maude et al., 2014), HIV (Kirby et al., 2017) or cholera (Nazia et al., 2018). However, it is also becoming a widely

implemented approach for non-infectious diseases such as Asthma (V. F. Jones et al., 2004), cancer (Sahar et al., 2019) or cardiovascular diseases (Mena et al., 2018). Infectious (communicable) and non-infectious (non-communicable) disease spatial epidemiological methods often share some common objectives, such as assessing the spatial distribution of disease and identifying the health impacts of environmental risk factors that can be used to generate public health action plans (Blangiardo, Boulieri, et al., 2020).

Methodologically, infectious disease modelling methods primarily focus on identifying the spatial or spatiotemporal clusters of infectious agents, how the disease spreads to the surrounding areas and the factors influencing spreading patterns as infectious diseases generally spread from human to human via air water or vectors (Rytkönen, 2004). In contrast, non-infectious disease modelling methods primarily focus on detecting whether certain environmental exposures change disease outcomes in an area over space or space-time (Blangiardo, Boulieri, et al., 2020). The statistical modelling of spatial variations in the risk of non-infectious diseases aims to advance the etiological hypothesis (Elliott & Wartenberg, 2004). These methods are similar to infectious disease modelling in identifying low and high-risk areas and contributing to the process of causation that can be related to diseases (Kroll et al., 2015). A spatial approach to disease analysis relevant to non-infectious disease is linked to potential environmental and socioeconomic drivers and their spatial heterogeneity (Kroll et al., 2015). Environmental factors such as increased food availability, availability of green space, air pollution, and walkability can often be linked to non-infectious diseases and fitted into the statistical modelling structure (Blangiardo, Boulieri, et al., 2020). In contrast, infectious disease modelling the infectious agents such as proximity to other humans (if human to human

infections), habitats of infectious disease agents, or socioeconomic inequalities may be fitted into the modelling structure (C.-H. Lin & Wen, 2022).

Some of the most common methods to determine and analyze non-infectious disease clusters are the Direction method, Kulldorff's spatiotemporal scan statistic, Jacquez's k-nearest neighbour method, Grimson's method, Knox's Method and Mantel's method, which have also been used for infectious disease analysis (Blangiardo, Boulieri, et al., 2020; Nagy & Negru, 2014). Kulldorff's scan statistic methods have been used to detect spatial and spatiotemporal clusters for both infectious and non-infectious diseases (Blangiardo, Boulieri, et al., 2020; Kroll et al., 2015). For instance, it has been used to identify clusters of infectious diseases such as tuberculosis (Onozuka & Hagihara, 2007) or dengue fever (Hohl et al., 2016) as well as for non-infectious diseases such as childhood cancer (Kreis et al., 2019) or asthma (Souza et al., 2019). Meanwhile, a further development has been the detection of spatial variations in temporal trends or the SVTT method for non-infectious disease modelling that extended the traditional scan statistics to estimate the time trend using a regression-based model (Moraga & Kulldorff, 2016). The quadratic SVTT method has been applied more specifically to cervical cancer data in women in the USA to find areas with unusual cervical cancer trends in white females using historical data (Moraga & Kulldorff, 2016).

Ecological niche modelling is a spatial approach that associates the occurrence location of infectious disease with a set of environmental variables allowing for the prediction of infection risk as an unknown location based on the environmental similarities (Z. Du et al., 2014; P. Jia et al., 2019). The ecological niche modelling outputs the prediction of geographic ranges for the infectious agents (P. Jia et al., 2019). In the non-infectious disease contexts, the location

of non-infectious disease cases or high prevalence of risk of non-infectious disease, based on existing evidence, could be considered as the presence of data to be linked with spatial data sets of lifestyles, behaviour, sociodemographic information and natural or build environments (P. Jia et al., 2019).

The Bayesian hierarchical models have also been used for both infectious and non-infectious disease modelling. In the context of non-infectious diseases, it generally includes demographic variables such as age, sex, ethnicity and socioeconomic status (Stringhini et al., 2018). A study by Goicoa et al. proposed a Bayesian age-space-time CAR model to study prostate cancer, a non-infectious disease, across the 50 provinces in Spain for age groups over 25 years while accounting for all pair-wise interaction (Goicoa et al., 2016). The study used a ranking of all provinces according to to identify the high-risk groups. The Bayesian Age-Period-Cohort modelling and prediction (BAMP) model is another method that provides a method of analyzing incidence and mortality data of non-communicable diseases, i.e. lung cancer (Y. Jiang et al., 2022) using a Bayesian version of age-period-cohort model (Schmid & Held, 2007). A hierarchical model is assumed with a binomial model in the first stage, and smoothing priors for the age, period and cohort parameters are further selected (Havulinna, 2011).

Shared component models were originally developed for two diseases, including common components such as common risk factors and a disease-specific one which can point towards specific risk factors, otherwise may be masked for a single disease modelling (Ibáñez-Beroiz et al., 2011). Shared component modelling has been widely used for non-infectious diseases such as cancer research. For example, shared component modelling was applied to male and female lung cancer and later extended to jointly model multiple diseases on the oral cavity,

esophagus, larynx and lung cancers in males in the 544 districts of Germany from 1986 to 1990 (Knorr-Held & Best, 2001). Recently, it was further extended to jointly model brain cancer incidence and mortality using Bayesian age and gender-specific shared component models (Etxeberria et al., 2018).

BaySTDetect, developed by Li et al., is a Bayesian detection method for short-time series of small area data between two competing space-time models that is primarily applied to detect unusual trends of non-infectious diseases (G. Li et al., 2012). The method was applied to detect unusual trends such as asthma and chronic obstructive pulmonary disease in England across mortality, hospital admissions and general practice drug prescription data (G. Li et al., 2012). FlexDetect and STmix are some simulation-based mixture models that have been primarily developed for non-infectious disease modelling. The STmix can identify areas where anomalies are present but not the temporal points, while FlexDetect, which is computationally more intensive, uses mixture modelling (Blangiardo, Boulieri, et al., 2020). More recently, another prediction model was developed for the management of non-infectious diseases and applied among older Syrian refugees during the COVID-19 pandemic in Lebanon. The outcome variable was the self-reported inability to manage any non-infectious disease, which included the following conditions such as cardiovascular disease, chronic respiratory disease, diabetes or hypertension, and predictor variables such as age, sex, residence, education, smoking status and food security was used (McCall et al., 2022).

The research methods of non-infectious diseases generally focus on detecting trends, highlighting unusual changes and consequently assisting in outlining the emerging non-infectious disease risk factors using statistical analysis (Robertson et al., 2010). Non-infectious

disease research designs share common objectives with infectious diseases, such as generating information to guide public health actions and detecting the health impact of environmental exposures or environmentally driven disease vectors (Blangiardo, Boulieri, et al., 2020; Carroll et al., 2014). However, methods for non-infectious have largely been based on detecting whether the outcome of interest shows a particular behaviour in an area, period or time or in space-time compared to the entire study region (Blangiardo, Boulieri, et al., 2020). While the research methods and the spatial statistical modelling of infectious versus non-infectious show some similarities and some differences, the choice of appropriate methods to model a disease ultimately depends on various factors that may influence disease processes and, most importantly, on the objectives of the study, data characteristics and computational resources (Graham et al., 2004).

2.3. Methods used in the three manuscripts

Infectious diseases can spread over space and time through susceptible populations of individuals by geographical processes (Kuebart & Stabler, 2020). Infectious Disease dispersion can be by expansion or relocation, and the spread process is generally mapped by concentrating on the increasing number of cases and then comparing across the neighbourhoods (Schærström, 2009). The spatiotemporal spread can be mapped by focusing on the changes in the total number of cases at a space at different points in time (Gianquintieri et al., 2020). The route of transmission or transmission path can be determined by proximity that spreads to the nearby area or by hierarchical levels, spreading stepwise between places.

Disease mapping is one of the most common ways to visualize the dynamics of an infectious disease in order to easily identify spatial dynamics and select areas of interest for

future studies. Generally, infectious disease outbreaks are defined by the increased number of cases in a local area (D. Chen, 2014). However, the movements of infectious disease outbreaks into surrounding areas are a spatiotemporal process (Niu et al., 2020). Infectious diseases can show heterogeneity in cases and mortality rates, while the process of contagion can disperse from the source to the neighbouring areas (Cordes & Castro, 2020). During an epidemic, it is important to display the outbreaks as they are happening and possible spread to the nearby neighbourhoods to provide deeper insights into understanding the spatial epidemiology of the disease that can allow a more effective response to outbreaks.

The first example of visualizing outbreaks is recorded from 1854 when Dr. John Snow plotted the location of water pumps in London during a cholera outbreak and identified the source water pump. More recently, Reyes et al. observed the spatiotemporal spread patterns of the 1918 influenza pandemic in British India by using using a travel network and a likelihood approach to predict the patterns of pandemic spatial spread. The output map shows the travel network connectivity between districts using nodes across British India, and the nodes with the high excess mortalities are marked in red to explain the spatial spread (Reyes et al., 2018). Another study by daCosta et al. visually examines the spatiotemporal diffusion of influenza A (H1N1), identifies the starting point of an epidemic, and further uses multiple maps of the time of onset and the correlation between effective distance trees across different spatial units (da Costa et al., 2018). A study by Cooper et al. used telehealth data to perform a spatiotemporal spread analysis of COVID-19. This study uses SaTScan to create high-risk cluster maps from multiple dates in the UK and then visually compares the spread using multiple dates (Cooper et al., 2008). Jeefoo et al. mapped the adjusted incidence per 1,000 inhabitants and per year from

1999 to 2007 (8 maps) analyze and compared visually the global spatial spread patterns of Dengue disease in Thailand (Jeefoo et al., 2011). These studies modeling infectious diseases, however, map does not include the direction of the spread or take temporal structures into account.

Some work has also been done to visualize various epidemic outbreaks, including COVID-19. For instance, Pourhasemi et al. have shown the temporal trend of the COVID-19 first outbreaks in Iran for each geographic unit using a choropleth map (Pourghasemi et al., 2020). Another study in Lombardy, Italy, used Emergency medical services parameters to identify the beginning of trends from emergency calls and EMS ambulances dispatches using a choropleth map to understand the spatiotemporal diffusion trends of COVID-19 (Gianquintieri et al., 2020). Long et al. used a spatiotemporal model that used a multifractal scaling method to show the spatial spread pattern of COVID-19 in China (Y. Long et al., 2020). However, these spatial transmission models do not show the origin, direction, and magnitude of the transmission processes.

Mapping COVID-19 spatial spreads primarily consist of comparing maps of COVID-19 new or cumulative cases from multiple time points in a spatial area to compare the outbreak progression visually (Cuadros et al., 2020; Y. Feng et al., 2020; Gao et al., 2021; H. Yu et al., 2021). For example, Cuadros et al. mapped the cumulative number of COVID-19 cases estimated from the spatial adjusted and non-spatial adjusted models on March 20, April 20 and May 10 of 2020 in Ohio, US, at the country level to visually compare the spatiotemporal spread patterns in these three months across the counties (Cuadros et al., 2020). Similarly, Feng et al. used local Moran's *I* statistic method to create ten high and low-risk cluster maps between 17

January 2020 to 20 march 2020 and visually compared these ten maps to identify the progression of disease across the cities in China (Y. Feng et al., 2020).

Existing spatial transmission or spread models often overlook the temporal structures, the direction and the magnitude of the transmission processes within the surrounding neighbourhood networks and primarily compare disease hotspot maps from multiple time points to visualize the dispersion processes. Since the transmission process of the disease may vary geographically and in time, it is important to investigate the transmission dynamics of the diseases using a novel model at a local scale. These types of maps can be effective in locating the origin of an outbreak and depicting outbreak progression over time.

In chapter 5, a novel approach was adopted for visualizing the spatial spread of the COVID-19 outbreaks at a neighbourhood level, possibly suggesting local transmission patterns and the direction of movements of a highly infectious disease in an urban context. In Chapter 5, these spatiotemporal processes as well as the direction of the diseases within the neighbourhood networks, are captured with the use of nearest neighbourhood networks with assumptions of two weeks as a temporal incubation period and defining an outbreak as the temporal cases larger than the study period temporal mean. Previous studies have emphasized the relevance of networks of infections and places of infections (M. Li et al., 2022; Osei et al., 2011; Suryowati et al., 2018). The neighbourhood network is related to the hypothesis that states that in urban spaces, infected individuals produce a transmission network in their households and in close neighbourhoods that will enhance the spread of the disease to new individuals in the same area or adjacent neighbourhoods (Perles et al., 2021). This study extends past work by providing a novel visualizing technique that incorporates the neighbourhood networks and temporal structures for

mapping the spatiotemporal dynamics of infectious disease outbreaks in one map and identifies the source and dispersion neighbourhoods, magnitude and direction of the outbreaks. The study also assesses the sources of infections in dispersed neighbourhoods and identifies the primary sources of infections in dispersed neighbourhoods.

A limitation of this method includes the reliability of the definition of the outbreak and the incubation periods to appropriately model the spatiotemporal dynamics of outbreaks. To address this limitation, sensitivity analyses were further performed to observe whether the changes in the definition of the outbreaks or the 2-week incubation period demonstrate significant variations in the spatiotemporal spread pattern.

This dissertation analyzes the spatial heterogeneity of the risk for COVID-19 at the neighbourhood level in Toronto, Canada, in the three manuscripts (Chapters 3-5). While traditional spatial statistical methods for identifying infectious disease hotspots use a frequentist approach which is more common, others use Bayesian approaches (Law et al., 2015). Each of the three manuscript chapters in this dissertation uses different methods to identify high-risk areas. For example, the widely used test-based method SaTScan (frequentist) in Chapter 3, Bayesian spatiotemporal models to identify the area-specific relative risk in Chapter 4 and a model that uses disease persistence and incidence rate to identify hotspots in Chapter 5. Hotspot prioritization using outbreak duration or disease persistence has been previously used for other infectious diseases (Debes et al., 2021; Ngwa et al., 2021) but has not been explored in the case of COVID-19.

Partaking different methods in the three manuscripts also allowed comparison among the methods by exploring whether they produce comparable results and the degree of these

similarities and dissimilarities. In Chapter 6, these differences and similarities are analyzed by using Cohen's Kappa Coefficients test to evaluate the degree of concordance (similarities) and discordance (dissimilarities) among the geographic patterns of high and low-risk neighbourhoods across the three studies. Additionally, the benefits and limitations of each of these hotspot methods are discussed, and recommendations are provided.

The accuracy of the remotely sensed LST data may depend on multiple parameters such as atmospheric effects, viewing angle and surface parameters. Remote sensing studies in the past have identified that the LST retrieval process using Land Surface Emissivity (LSE) model by Sobrino et al. (2008) (Sobrino et al., 2008), with an atmospheric correction provides a high estimation accuracy (Z.-L. Li et al., 2013; Sobrino et al., 2008; X. Yu et al., 2014; Zhao et al., 2009). In Chapter 4 this method was adopted while correcting the images using atmospheric correction parameters collected by the National Aeronautics and Space Administration (NASA) (NASA, 2021) to retrieve LST at the neighbourhood scale. To the best of my knowledge, no previous studies on COVID-19 have adopted correction methods in the calculation of LST. Previous regression studies that have modelled the influence of LST variations in COVID-19 studies have also only used LST data from one or two temporal dates and have used a larger spatial unit of analysis (i.e. city, county) which may not be realistic representation of the surface temperature or the spatiotemporal fluctuations (Das et al., 2020; Hassan et al., 2021). The methodological framework of this dissertation improves upon existing literature by using a higher spatial (neighbourhood-scale) and temporal (79 weeks) resolution satellite imageries in Chapter 4, which is more spatially representative and may have provided a more accurate estimate for using the atmospherically corrected data on LST.

Existing spatiotemporal COVID-19 studies have predominantly used testing-based methods to classify and identify hotspots (Nazia et al., 2022). However, in order to identify the spatiotemporal trends, these test-based methods, such as SaTScan, cannot be assessed for a specific small area and are stable or show increasing or decreasing patterns (Sherman et al., 2014). Earlier COVID-19 spatial studies have not assessed the spatiotemporal trends of COVID-19 in a study area (Franch-Pardo et al., 2021; Nazia et al., 2022). In Chapter 4, in order to identify the spatiotemporal and temporal trends, Bayesian spatiotemporal models with a set of observed covariates and unobserved spatial random effects, temporal random effects and three different types of space-time random effects are explored (Knorr-Held and Besag, 1988; Knorr-Held 2000). Chapter 4 advances past work by modelling the dynamics of the temporal and the increasing, decreasing or stable spatiotemporal trends of COVID-19 at the neighbourhood scale in Toronto using the Bayesian approach, which, to the best of my knowledge, has not yet been modelled for COVID-19. The Bayesian approach also allows dealing with uncertainties related to the data, the process and model parameters (Haining & Li, 2020) since it has the capacity to account for missing data, measurement errors and ecological bias (Knorr-Held & Raßer, 2000; Law et al., 2014).

In chapter 3, the widely used spatial regression models (Spatial Error Model and Spatial Lag model) and in Chapter 4, the Bayesian hierarchical models are applied to identify the important socioeconomic risk factors to influence the overall COVID-19 risk for Toronto. The findings of the important socioeconomic risk factors from these two methods are further discussed in Chapter 6.

Chapter 3 employs geographically weighted regression models (GWR) that overcome the limitation of global regression models by taking spatial heterogeneity into consideration while calculating the spatial interaction between the outcome and explanatory socioeconomic variables in each neighbourhood (C. H. Lin & Wen, 2011; Maiti et al., 2021). In particular, the new version of GWR, termed multi-scale geographically weighted regression (MGWR), is applied, which is comparable to Bayesian spatially varying coefficients (SVC) models. Compared to the classical GWR models that operate at the same spatial scale, the MGWR model potentially provides a more flexible and scalable framework for examining the multi-scale process (Fotheringham et al., 2017). A recent study by Fotheringham et al. compared the GWR model and MGWR model, where the MGWR model provide more valuable information on the scale at which different processes operate (Fotheringham et al., 2017). This study by Fotheringham et al. also compared MGWR to the Bayesian (SVC) models (Gelfand et al., 2003) and identified that the MGWR model provides an alternative to the SVC model that is arguably computationally less demanding (Finley, 2011), scales more easily and, to some, is more intuitive (Fotheringham et al., 2017).

A common issue with the traditional GWR model is the multiple testing issues leading to an excess of false positives that may make the significance of the local parameter estimates questionable (A. R. da Silva & Fotheringham, 2016). However, this issue is hardly taken into account in past COVID-19 studies that adopted the GWR methods. To solve this issue, in Chapter 3, a newly developed correction method by da Silva and Fotheringham (A. R. da Silva & Fotheringham, 2016) is applied to the MGWR model for an effective correction to the significance level to account for the multiple testing issues.

2.4. Spatial and temporal Structures

While infectious disease risk may be known to vary within space, the temporal dynamics of the disease risk also may vary locally (López-Quílez, 2019). The fundamental properties of spatial and spatiotemporal data are dependence and heterogeneity (B. Jiang, 2015; Robertson et al., 2010). For spatial and spatiotemporal data, values that are close in space or in space-time are likely to be dependent (Robertson et al., 2010). In spatiotemporal data, health outcome observed in an area during an interval of time generally contains some information about the health outcome in nearby areas within the same time. In Chapter 3, the geographies of COVID-19 risk were analyzed by considering space and time as a lens through which we observed the outbreaks as a process and the spatially varying regression coefficients were captured within each spatial unit. In Chapter 4, four Bayesian spatiotemporal models were fitted with one or more sets of random effects parameters. The first model uses spatial and temporal random effects terms, while model 2 to 4 adds space-time interaction effects to allow space-time inseparability.

2.5. Scale, aggregation and areal unit of analysis

This study uses neighbourhood scale as the areal unit of analysis. A major challenge in mapping infectious diseases, particularly in Canada, is the lack of access to individually geocoded or disaggregated data to preserve participants' privacy and confidentiality (Beenstock & Felsenstein, 2021). Due to privacy concerns, Statistics Canada also suppresses the socioeconomic and demographical information for the residents. Disease case information for each individual is generally aggregated to a spatial administrative unit before they are released to the public. The use of spatially aggregated to large areas to preserve individuals' privacy can sometimes lead to a modifiable areal unit problem (MAUP) which can challenge researchers to

fully reflect the reliability of the results (Byun et al., 2021). MAUP refers that aggregating these data to different sizes or geographical units for spatial analysis can cause in terms of accuracy, quality, possible bias and confounding factors, particularly for a large areal unit of analysis (Manley, 2014). One of the best methods to resolve the MAUP issues in aggregated studies is to use data aggregated to the smallest administrative units available (high granularity). In Toronto, the neighbourhood scale is the smallest areal administrative unit available for COVID-19 spatial datasets that is a high-resolution geographically referenced data. This spatial scale offers an opportunity to facilitate a small-area analysis of local geographic variations in diseases for case, socio-economic and environmental datasets; therefore, the neighbourhood-scale was chosen as the unit of analysis for this dissertation.

2.6. Toronto as a case study

The city of Toronto covers a large geographic area with 140 distinct neighbourhoods that allow us to observe COVID-19 heterogeneity at a small local level (Fig. 3.1). This thesis uses the urban city of Toronto as a case study across all three manuscripts as Toronto serves as an ideal geographical area for an ecological study of COVID-19 using spatial and spatiotemporal analysis of COVID-19. Toronto, the capital of Ontario, is the most densely populated city in Canada (Statistics Canada, 2017), with over 2.9 million people living in a 630 square kilometre land area (City of Toronto, 2020a) and has been severely impacted by COVID-19. In Canada, the first two cases of COVID-19 in Canada were reported in Toronto on January 21, 2020. The COVID-19 case data are also released weekly and the full dataset are publicly available, allowing a spatiotemporal analysis of the study area.

The availability of the spatiotemporal datasets, the large population of the City of Toronto and high case counts can provide a suitable sample size to ensure the reliability and stability of the results of a small-area analysis of an infectious disease. Additionally, the city of Toronto has a diverse population with socioeconomic variances and, therefore, can effectively capture the effects of these variances on COVID-19 incidence. Therefore, Toronto serves as an ideal case study to represent urban areas for the primary goal of this dissertation which is to understand the spatial epidemiology of COVID-19 in urban neighbourhoods. A limitation of this dataset may include unreported cases, asymptomatic cases, cases missing neighbourhood information or the availability of relevant latest datasets such as mobility and air pollution at the neighbourhood-level. However, these missing/unreported cases may be spatially random, and some of this missingness are also accounted for with the use of Bayesian modelling techniques in this dissertation.

Overall, this case study provides a methodological framework for a detailed study of infectious disease epidemiology, which can serve as an example for other research studies in similar settings for COVID-19 or future epidemic modelling of urban infectious disease burdens with the availability of spatial and temporal data.

2.7. Datasets

This dissertation uses open-source datasets that are publicly available with three primary types of data: (1) Toronto neighbourhood profiles that contain census data with neighbourhood-level population characteristics and boundary files, (2) COVID-19 cases, and (3) remotely sensed datasets (Land surface temperature data), acquired and calculated from NASA Landsat series of Earth Observation satellite images (*USGS- Landsat*, 2021). More detailed descriptions of the data are described in Chapters 3-5.

2.7.1. Neighbourhood Profiles Dataset

The 140 neighbourhoods were defined based on Statistics Canada census tracts for the purposes of statistical reporting (Statistics Canada, 2021). The digital data of the geographic boundaries for these neighbourhoods were acquired from the open data portal of the city of Toronto (City of Toronto, 2020b). In Chapters 3 and 4, variables measuring the social, economic and demographic characteristics of the neighbourhoods were originally retrieved using the census 2016 population dataset collected and released by Statistics Canada (City of Toronto, 2017). The full descriptions of the variables are provided in Table 4.1.

2.7.2. COVID-19 case data

The epidemiological data of COVID-19 cases were originally collected by Toronto Public Health and extracted from the provincial Case & Contact Management System (CCM) by the city of Toronto. The case dataset contains the demographic, geographic, and severity information for all confirmed and probable, sporadic, and outbreak-associated cases. These data are updated on a weekly basis. More information on these COVID-19 case datasets is published by Toronto Public Health on https://open.toronto.ca/dataset/covid-19-cases-in-toronto/.

2.7.3. LST Data

In chapter 4, Landsat imageries (Landsat 7 and Landsat 8) were retrieved, and an atmospheric correction was performed to calculate the weekly average land surface temperature (LST) at the neighbourhood level for 89 consecutive weeks. Mostly Landsat 8 imageries (75) were used, and

if the imageries for a particular week were not available in the Landsat 8, Landsat 7 imageries were used instead. More details for the LST data are included in Chapter 4.

Nevertheless, these census and case datasets for the City of Toronto and the community-level LST data have not yet been fully exploited for COVID-19 assessment but could advance the understanding of community-level COVID-19 epidemiology in an urban environment and benefit COVID-19 intervention plans.

Chapter 3 Manuscript 1: Spatial and spatiotemporal clusters and the socioeconomic determinants of COVID-19 in Toronto neighbourhoods, Canada: An Observational Study[1]

3.1. Summary

The aim of this study is to identify spatiotemporal clusters and the socioeconomic drivers of COVID-19 in Toronto. Geographical, epidemiological, and socioeconomic data from the 140 neighbourhoods in Toronto were used in this study. We used local and global Moran's *I* and space-time scan statistic to identify spatial and spatiotemporal clusters of COVID-19. We also used global (spatial regression models), and local geographically weighted regression (GWR), and Multi-scale Geographically weighted regression (MGWR) models to identify the globally and locally varying socioeconomic drivers of COVID-19. The global regression model identified a lower percentage of educated people and a higher percentage of immigrants in the neighbourhoods as significant predictors of COVID-19. MGWR shows the best-fit model to explain the variables affecting COVID-19. The findings imply that a single intervention package for the entire area would not be an effective strategy for controlling the disease; a locally adaptable intervention package would be beneficial.

[1] The citation for this article is Nushrat Nazia, Jane Law, Zahid Ahmad Butt, Spatiotemporal clusters and the socioeconomic determinants of COVID-19 in Toronto neighbourhoods, Canada, Spatial and Spatiotemporal Epidemiology, 2022, 100534, ISSN 1877-5845, https://doi.org/10.1016/j.sste.2022.100534.

3.2. Introduction

In Canada, the coronavirus disease (COVID-19) burden is unevenly distributed spatially, while the highest number of cases are witnessed in Toronto (*COVID-19 Tracker Canada*, 2021; Detsky & Bogoch, 2020). COVID-19 is a part of a family of enveloped single-strained RNA viruses that can cause acute and chronic communicable respiratory diseases in humans (World Health Organization, 2020). Toronto, the most densely populated city in Canada, has been severely impacted by COVID-19 and has become an epicentre of COVID-19 outbreaks (City of Toronto, 2021a). COVID-19 was first reported in Wuhan, China, in December 2019 (Kang et al., 2020; Shereen et al., 2020). As of March 7, 2022, the large-scale outbreaks of COVID-19 have contributed to over 440 million cases and 5.9 million deaths worldwide (WHO, 2021). The World Health Organization (WHO) declared the COVID-19 outbreak a global pandemic on March 11, 2020 (Cucinotta & Vanelli, 2020). The first two cases of COVID-19 in Canada were reported in Toronto on January 21, 2020, from a couple who had recently returned from Wuhan, China. The local government of Toronto has adopted several control strategies that include emergency lockdowns, stay-at-home orders, increased testing, contact tracing capacities, and closure of in-person schools and non-essential businesses. Despite these ongoing measures, Toronto continues to experience a rise in cases, creating an enormous challenge for public health as well as causing economic and social burdens.

Spatial studies in COVID-19 showed wide variances in the distribution of case and mortality rates among different communities across space. The lower socioeconomic groups have historically been shown to have a disadvantage in the diagnosis and mortality rates of infectious diseases. The city of Toronto has a diverse population of 2.99 million (2020) with

varying socioeconomic statuses across the neighbourhoods (City of Toronto, 2021a). People with low socioeconomic or marginalized status, such as minorities and low-income individuals, might be forced to leave their homes to maintain income or live in congregate settings, which places them and their neighbourhoods at a higher risk during this pandemic (Cordes & Castro, 2020; Y. Sun et al., 2020a; Vaz, 2021). Identifying disease clusters and understanding the driving factors for these clusters using spatial analytical approaches can provide us with a more realistic view of the issue in Toronto compared to the traditional simple maps (Cordes & Castro, 2020).

Geographic Information Systems and spatial analysis have been established as important tools in infectious disease surveillance. Proximity is an important factor in infectious disease distribution and diffusion processes. Spatial analysis is based on Tobler's first law of geography, stating that locations that are closer have more similar attributes than locations that are further apart (H. J. Miller, 2004). Infectious diseases can show heterogeneity in cases and mortality rates, while the process of contagion can diffuse from the source to the neighbouring areas (Cordes & Castro, 2020). Spatial analytical methods using community-based datasets are particularly important for new emerging diseases such as COVID-19 (Ali et al., 2009) to understand disease etiology and the transmission process in the communities (Chowell & Rothenberg, 2018b). The results can help slow disease transmission rates in the study area.

In the case of COVID-19, previous spatial studies used space-scan statistic (Acharya et al., 2020; Alkhamis et al., 2020; Andersen et al., 2021b; Andrade et al., 2020; Arashi et al., 2020; Azmach et al., 2020; Ballesteros et al., 2020; Benita et al., 2020; Chakraborty, 2021; Chow et al., 2020; Cordes & Castro, 2020; Da Silveira Moreira, 2020; Desjardins et al., 2020; M. C. Ferreira,

2020; R. V. Ferreira et al., 2020; Gomes et al., 2020; Greene et al., 2020; Han et al., 2021; Hohl et al., 2020; Islam et al., 2021; S. Kim & Castro, 2020; Ladoy et al., 2021; Leal-Neto et al., 2020; Masrur et al., 2020) and local Moran's *I* (Bilal et al., 2020; Han et al., 2021; H. Li, Li, et al., 2020; Saffary et al., 2020; Xie et al., 2020; Xiong et al., 2020; L. Ye & Hu, 2020), MST-DBSCAN (De Ridder et al., 2021; Ridder et al., 2020), GeoMEDD (A. Curtis et al., 2020) to locate clusters of increased risk of COVID-19. A few studies have used global spatial regression models (Cao et al., 2020; Demenech et al., 2020; Feinhandler et al., 2020; Sannigrahi, Pilla, Basu, Basu, et al., 2020b; Y. Sun et al., 2020b; You et al., 2020), Geographically Weighted Regression (GWR) (Das et al., 2020; Han et al., 2021; J. Huang et al., 2020; Islam et al., 2021; Iyanda et al., 2020; Mollalo et al., 2020; Sannigrahi, Pilla, Basu, Basu, et al., 2020b; Shariati, Jahangiri-rad, et al., 2020; Snyder & Parks, 2020) and MGWR (Multiscale Geographically Weighted Regression) (J. Ma et al., 2022; Maiti et al., 2020; Mansour et al., 2021; Middya & Roy, 2021) to understand the contributing factors that may influence the risk of COVID-19. Some studies have linked socioeconomic influences to explain the variations in COVID-19 incidences or mortality rates. Some of the key factors to have an influence on the incidence or mortality rates of COVID-19 include income (Abedi et al., 2020; Chaudhry et al., 2020; Cordes & Castro, 2020; Maiti et al., 2020; Sannigrahi, Pilla, Basu, Basu, et al., 2020a), poverty rates (Y. Chen & Jiao, 2020a; Fielding-Miller et al., 2020; Goutte et al., 2020; Richmond et al., 2020; Sannigrahi, Pilla, Basu, Basu, et al., 2020a), education (Abedi et al., 2020; Cordes & Castro, 2020; Goutte et al., 2020; Y. Wu et al., 2020) ethnicity or minority status (Andersen et al., 2021a; Y. Chen & Jiao, 2020a; Cordes & Castro, 2020; Kathe & Wani, 2020; Maiti et al., 2020; Niedzwiedz et al., 2020; F. Sun et al., 2020a; Y. Sun et al., 2020a), immigrant population

(Borjas, 2020; Fielding-Miller et al., 2020), and unemployment rate (Goutte et al., 2020; Y. Sun et al., 2020a).

Even though Toronto remains at heightened risk in this pandemic, spatial studies on the burden of COVID-19 in Toronto have so far been very limited. A detailed assessment is critical to identify the hotspots and the key spatial drivers for a more efficient intervention plan. This study is to fill the gaps of earlier studies by performing a comprehensive assessment of spatial dynamics of the COVID-19 outbreak in Toronto at the neighbourhood level. The main goal of this study is to improve our current understanding of the disease hotspots and provide information on the spatial association between the socioeconomic factors and COVID-19 outbreak in order for infection prevention and mitigation. The primary objective of this study is to identify the spatial and spatiotemporal clusters of COVID-19 incidences in Toronto, and the secondary objective is to identify globally and locally variable socioeconomic drivers of the COVID-19 incidences.

3.3. Methods

3.3.1. Study area, study population and data

Toronto is the capital city of Ontario and the fourth largest city in North America. The city, located on the Southwestern shores of Lake Ontario (Figure 3.1), has a population density of 4,692 people per square kilometre, making it the most densely populated city in Canada. The city of Toronto consists of four community council areas: Etobicoke York, North York, Toronto, and East York and Scarborough (Figure 3.1). Toronto has a very diverse population, with a 51.2% of Toronto's population being a visible minority, and 51.2% of the population are immigrants (born outside of Canada) (City of Toronto, 2021a). Toronto, based on the 2016 census, comprises

immigrants from Asian (53.4%), European (23.6%), Americas (16.8%) and Africa (6.1%) origins. 49% of the Asian immigrants have immigrated from Chinese and the Philippines, and 11.7% of Asian immigrants are from India. 37.5% of the American immigrants are of Jamaican and Guyanese origin. 41.7% of the European immigrants are from Italy, Portuguese, and United Kingdom. 55% of the African immigrants are from Ethiopia, Egypt, Kenya, Nigeria, Somalia and South Africa. The city has an unemployment rate of 6.4%, and 20.2% population lives in a low-income bracket. The direct and indirect impact of this pandemic's burdens can pose significant threats for a densely and diversely populated city like Toronto.

The study area is comprised of 140 geographically distinct neighbourhoods in Toronto (Figure 3.1). The neighbourhoods were defined based on Statistics Canada census tracts for the purposes of statistical reporting (Statistics Canada, 2021). The140 neighbourhood profiles in Toronto contain social, economic, and demographic details using the census 2016 population dataset collected and released by Statistics Canada. Our study used a population size of 2.7 million (2016 census dataset) for analysis. Further details of these neighbourhood profiles datasets can be found on the city of Toronto website (Neighbourhood Profiles – City of Toronto) (City of Toronto, 2020b). The population dataset at the neighbourhood level was obtained from an open-source (https://open.toronto.ca/dataset/wellbeing-toronto-demographics).

The epidemiological data were collected by Toronto Public Health that contained the geographic and demographic details for all confirmed and probable cases of COVID-19. Confirmed cases are defined by a person with confirmation of SARS-CoV-2 infection documented by detection of at least one specific gene target by a validated laboratory-based nucleic acid amplification test (NAAT) assay (e.g. real-time PCR or nucleic acid sequencing)

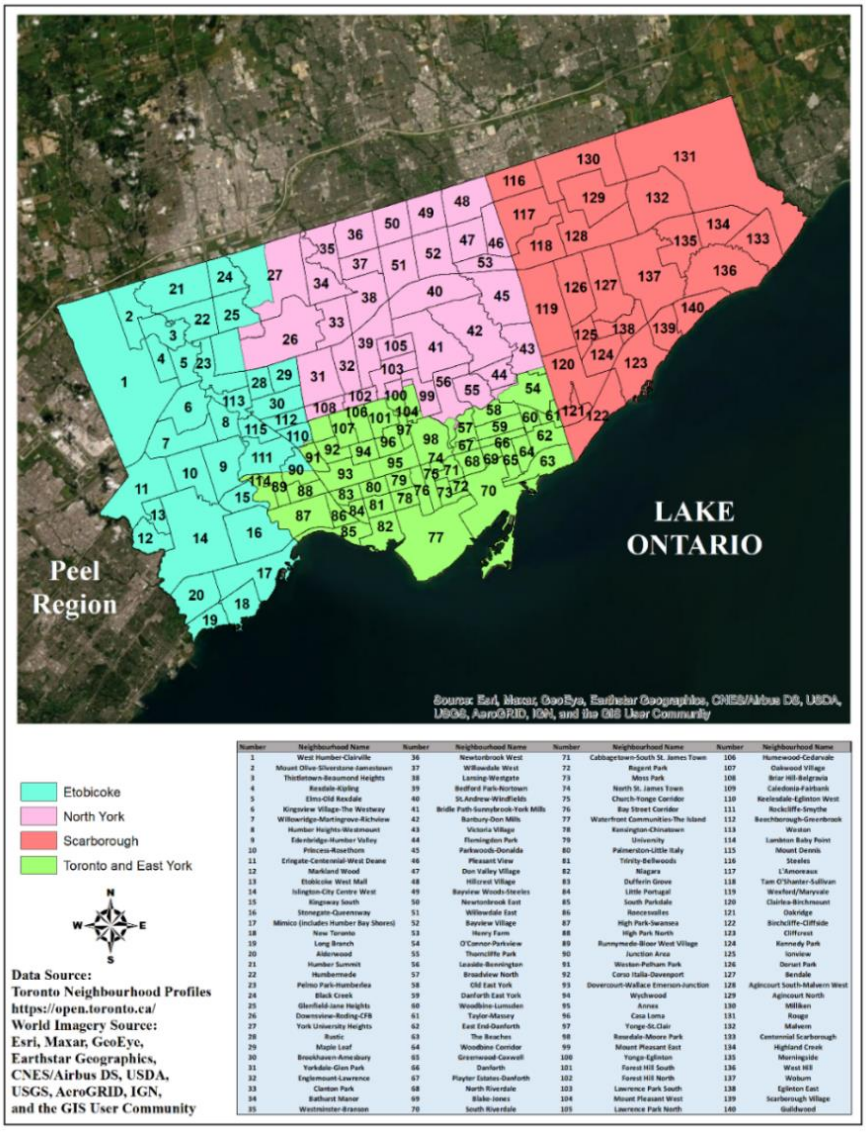
performed at a community, hospital or reference laboratory (e.g. Public Health Ontario Laboratory or National Microbiology Laboratory) or a validated POC NAAT has been deemed acceptable by the Ontario Ministry of Health to provide a final result or demonstrated diagnostic rise with a 4-week interval in viral-specific antibody in serum or plasma using a validated laboratory-based serological assay for SARS-CoV-2. The probable cases are defined by a person with symptoms compatible with COVID-19 and had high-risk exposure with a confirmed case of COVID-19 or was exposed to a known cluster or outbreak and in whom a laboratory-based nucleic acid amplification test (NAAT-based assay (e.g. real-time PCR or nucleic acid sequencing) for SARS-CoV-2 are inconclusive or has not been completed or had SARS-CoV-2 antibody detected in a single serum, plasma, or whole blood sample using a validated laboratory-based serological assay for SARS-CoV-2 collected within four weeks of symptom onset or had a POC NAAT or POC antigen test for SARS-CoV-2 completed and the result is presumptive positive or had a validated POC antigen test for SARS-CoV-2 completed and the result is positive (Canada, 2021b).

The case datasets are extracted from the provincial Case & Contact Management System (CCM). The epidemiological datasets are updated weekly (City of Toronto, 2020a). The case dataset includes the episode date and neighbourhood attributes of the infected individuals. For this study, the COVID-19 case data between January 1, 2020 – January 31, 2021, were extracted for analysis. The cases were plotted to the centroid of the infected individual's neighbourhood. During the study period, a total of 87,501 cases (71,940 sporadic cases and 15,561 outbreak-associated cases) of COVID-19 were diagnosed in Toronto, including the first reported case (January 21, 2020). Use of outbreak-associated cases, which are generally in healthcare (e.g.,

long-term care homes, hospitals) and residential or congregate settings (City of Toronto, 2020a), can potentially create a bias towards clustering in outbreak areas (17.71% of total cases). Therefore, this study excluded the outbreak-associated cases to control for potential bias and included only the community-based sporadic cases. Among the total sporadic cases, 1,185 (1.64%) cases were excluded due to missing neighbourhood information, leaving 70,755 sporadic cases for analysis.

We considered five socioeconomic covariates (2016 census) based on the literature review and available datasets. These covariates were previously identified as potential risk factors in some COVID-19 studies. The covariates included: i) percentage of immigrants (individuals who were born outside of Canada) (Government of Canada, 2017a), ii) percentage of the population aged 25-64 years with a lower level of education (not having a university certificate, diploma or a bachelor degree)[79], iii) prevalence of low income (living in a low-income household based on the low-income cut-off (LICO) table representing the poverty line) (Government of Canada, 2021a), and iv) unemployment rate (population over 15 years and unemployed) (Government of Canada, 2021c). While mortality and hospitalization rates may differ across different age groups, there is not enough evidence that the incidence rate of COVID-19 and transmission risks for sporadic cases vary significantly by sex or age (H. Li, Wang, et al., 2020). Since the entire population is at risk of contracting and transmitting the disease, we did not pursue any adjustment by age or sex.

Figure 3.1. The study area in Toronto, Ontario, Canada.



3.3.2. Spatial and spatiotemporal cluster analyses

3.3.2.1. The Global Spatial Autocorrelation

The global Moran's *I* by Anselin was first used in GeoDa, version 1.18.0, to assess whether the COVID-19 incidence rates (per 1000 population) in Toronto neighbourhoods display a tendency to cluster together and measure the extent of the correlation among neighbouring observations (Anselin, 2018). Global Moran's index is used to examine the absence or presence of spatial

autocorrelation in disease diffusion processes by comparing location and attribute similarities in the area. The value of the global Moran's *I* must show a clustering distribution pattern to find high or low-risk clusters for further analysis (Shariati, Mesgari, et al., 2020). The formula for calculating the global Moran's *I* index(Anselin, 2018) is shown in **Eq. 1**:

$$I = \frac{\Sigma_i \Sigma_j w_{ij} Z_i . Z_j / S_0}{\Sigma_i Z_i^2 / n} \quad (1)$$

where Z_i and Z_j represent the COVID-19 incidence rate variations in neighbourhood i and j, respectively, w_{ij} refers to the elements in the spatial weights matrix, neighbourhood i and j at study period, $S_0 = \Sigma_i \Sigma_j w_{ij}$, w_{ij}as the sum of all weights, and n represents the number of observations.

The value of global Moran's *I* can range between -1.0 and +1.0, where the positive value suggests the presence of a positive spatial correlation, while a negative value suggests a negative correlation. The higher the value of *I*, the stronger the spatial autocorrelation (S. Kim & Castro, 2020). Values close to 0 indicate no spatial autocorrelation and that the distribution of data is random (H. Li, Li, et al., 2020).

3.3.2.2. Spatial clustering by local Moran's *I* statistics

The local Moran's *I*, a local indicator of spatial association (LISA), was used in GeoDa, version 1.18.0, to evaluate the local level of spatial autocorrelation or dependency of spatial data and to visualize the possible high-risk or low-risk clusters (Anselin et al., 2010) based on COVID-19 incidences in different neighbourhoods across Toronto.

The formula (Anselin, 1995; H. Li, Li, et al., 2020) for calculating local Moran's index I_i is shown in **Eq. 2**:

$$I_i = x_i \sum_j w_{ij} x_j \quad (2)$$

where $x_i \; and \; x_j$ represents the COVID-19 incidence rates in neighbourhood i and j respectively, w_{ij} is the spatial weights matrix.

The global and local Moran's *I* tests were run using the first-order queen's contiguity spatial weights matrix that uses the values from all first-order neighbouring neighbourhoods in order to determine whether the area has a higher or lower mean assessing the degree of spatial autocorrelation. A permutation test was conducted using Monte Carlo simulations with 999 permutations to test the statistical significance of the clusters under the assumption that COVID-19 incidence rates are randomly distributed in the study area. The local Moran's *I* divide the neighbourhood polygons into four categories: high-high (hotspots), low-low (coldspots), high-low, and low-high, based on the type of spatial autocorrelations (Anselin, 1995). The high-high and low-low areas represent spatial clusters, and the high-low and low-high areas represent discordant patterns. The intensity value is calculated for each point, which shows the level of clustering of similar values around the point. The local Moran's *I* result showing local spatial autocorrelation of COVID-19 incidences in the Toronto neighbourhoods were presented in the form of cluster maps with a significance level of 0.1%, 1%, and 5%. Additionally, we have also performed the Bonferroni bound procedure to carry out an extensive sensitivity analysis to avoid the risk of obtaining false-positive results (Type I errors), and to check the robustness of the findings (not ultimately presented these results). We have presented the results that best represent the important High-High and Low-Low clusters.

3.3.2.3. The space-time scan statistic

Kulldorff's space-time scan statistic method was used in SaTScan™, version 9.7, to identify the space-time clusters of COVID-19 cases between January 2020 and January 2021 in Toronto neighbourhoods. SaTScan™ software is a widely used open-source spatial scan statistic software that utilizes Kulldorff's retrospective space-time permutation method to identify significant clusters in a study area (Kulldorff & Information Management Services Inc, 2009). SaTScan™ uses a moving cylinder with circular or elliptical windows across a study area that locates the spatial clusters that are significant during a specific period. A discrete Poisson probability model was chosen for the clustering analysis with the assumption that the disease cases have a Poisson distribution. The scan parameters with a time interval of one month ranged from January 1, 2020 – January 31, 2021. After a preliminary test, the spatial and temporal scanning windows were restricted to include 10% of the population at risk and 50% of the study period, respectively, to avoid a large cluster size. The clusters were tested for significance using 999 Monte Carlo simulations, and the clusters with a p-value<0.05 are considered to be significant high-risk clusters. The neighbourhoods within the significant high-risk clusters are identified as high-risk neighbourhoods. The relative risk of COVID-19 for a cluster is calculated using the ratio of observed to expected cases, comparing the risk within a cluster to the areas outside the cluster. The relative risk (RR) (Kulldorff & Information Management Services Inc, 2009; Rao et al., 2017) defined in **Eq. 3**:

$$RR = \frac{o/e}{(O-o)/(O-e)} \quad (3)$$

Where e is the expected number of cases in the cluster, o is the total number of observed cases within the cluster, and O is the total number of observed cases in the study area. The values of RR

for a cluster greater than 1 indicate a high COVID-19 incidence rate. A spatiotemporal map with the clusters and the relative risks (RR) of the neighbourhoods was created in ArcGIS version 10.8.1 to show the spatial variations of COVID-19 risks in the Toronto neighbourhoods.

3.3.3. Regression Analyses

We used five different global and local spatial regression models to understand the relationship between the socioeconomic variables and COVID-19 incidence rates. The models include three global regression models: ordinary least squares (OLS), spatial error model (SEM), spatial lag model (SLM), and two local regression models: geographically weighted regression (GWR) and multiscale GWR (MGWR). Before running these models, a bivariate regression analysis was conducted to select the explanatory variables. GeoDa version 1.18.0 was used for running the three global models. The local models were implemented in a stand-alone software MGWR version 2.2: Spatial Analysis Research Center (SPARC), Tempe, USA (https://sgsup.asu.edu/sparc/multiscale-gwr), developed by Fotheringham et al. (Fotheringham et al., 2017). ArcGIS version 10.8.1 was used for the mapping of all outputs.

3.3.3.1. Global Regression Models

The COVID-19 incidence rate per 1000 population was used as the dependent variable for the global models. Preliminary data analysis shows that the incidence rates were highly skewed, violating the normality assumptions of spatial regression models (D. Yu et al., 2010); thus, the log-transformed (base 10) incidence rates were used as the dependent variable in the models. The ordinary least squares (OLS) is a regression method investigating the relationship between the dependent and explanatory variables. OLS makes two major assumptions: the observations

are independent and constant across the study area, and there is no correlation between the error terms (Yandell & Anselin, 1990). A spatial error model (SEM) (Yandell & Anselin, 1990) is based on the assumption that there is a spatial dependence in the OLS model residuals generated from the OLS error term model. A spatial lag model (SLM) Spatial lag model (Yandell & Anselin, 1990) is based on a spatially-lagged dependent variable. The SLM model assumes dependency among the dependent and the independent variables. The SLM also assumes that an independent variable can depend on another independent variable in the neighbourhood region. The OLS model is expressed in **Eq**. (4) as:

$$y_i = \beta_0 + x_i\beta + \varepsilon_i \quad (4)$$

The mathematical expression for the SEM is shown in **Eq**. (5) as:

$$y_i = \beta_0 + \beta x_i + \lambda W_i\mu_j + \varepsilon_i \quad (5)$$

The SLM model is expressed in **Eq**. (6) as:

$$y_i = \beta_0 + x_i\beta + \rho W_i y_i + \varepsilon_i \quad (6)$$

Where yi is the COVID-19 incidence rate in neighbourhood *i*, x_i is the vector of the explanatory variable, ε_i is an error term, *Wi* is a vector of (nxn)spatial weights matrix, β is the vector of regression parameters, β_0 is the intercept, μ_i and μ_j are the error terms at neighbourhood *i* and *j*, respectively**,** λ is the coefficient of spatially correlated errors, and ρ is the spatial lag parameter.

However, in the case of COVID-19 in the Toronto neighbourhoods, as supported by SLM and SEM results shown later in the results, a spatial correlation exists between variables. Therefore, the interactions from the OLS are omitted from the results, and the spatial models (SEM and SLM) were considered better suited for this study. The AIC values were used in the

final model selection process to evaluate overall model accuracy and how well the model fits the data, and a lower AIC value indicates an improvement in model performance.

3.3.3.2. Local Regression Models (GWR & MGWR)

Two local models, GWR and MGWR, were applied to the same set of predictors used in the global models to explore the local spatial variation in the relationships with the COVID-19 incidence rates.

The GWR model is a local spatial regression model that makes assumptions that spatial interactions are non-stationary and that parameter estimates may spatially vary, which can not be explained by the global regression models (C. H. Lin & Wen, 2011). GWR takes spatial heterogeneity into consideration while calculating the spatial interaction among the dependent and explanatory variables and produces local regression parameter estimates at each observation location (C. H. Lin & Wen, 2011; Maiti et al., 2021). Geographically Weighted Regression (GWR) uses a local smoothing processing method to estimate the geographical functional form of regression coefficients non-parametrically (Nakaya, 2016). The GWR model is denoted in **Eq.** (8) as:

$$y_i = \sum_{j-0}^{m} \beta_j(\mu_j, v_j) X i_j + \epsilon_i \quad (8)$$

where at area i, y_i is the dependent variable (log of COVID-19 incidence rate), $\beta_j(\mu_j, v_j)$ is the jth coefficient, (μ_j, v_j) is the vector form of x, y coordinates, X_{ij} is the value of the jth explanatory parameter, and ϵ_i is the random error term (Iyanda & Osayomi, 2021).

However, the GWR models produce a single optimal bandwidth for all variables, which assumes that all factors affect COVID-19 rates at the same spatial scale (H. Yu et al., 2020). This assumption is given that different processes may affect COVID-19 rates at different spatial scales. This can result in an underestimation of the parameters, particularly in a large city such as Toronto, with a high population density (Leong & Yue, 2017). Therefore, we have also applied the MGWR model, which is an extension of GWR that allows for studying the relationship between variables at different scales (H. Yu et al., 2020). MGWR obtains a set of optimal covariate-specific bandwidths in which each bandwidth indicates the spatial scale at which a factor impacts the outcome variable (Fotheringham et al., 2017). The MGWR model can be formulated in **Eq**. (9) as:

$$y_i = \sum_{j-0}^{m} \beta_{bwj}(\mu_j, v_j) Xi_j + \epsilon_i \quad (9)$$

where $\beta_{bwj}$is the bandwidth used for calibration of the jth relationship (Iyanda & Osayomi, 2021), and the rest of the parameters are the same as **Eq.** (8).

Additionally, geographically weighted regression models generally ignore the multiple testing issues that can lead to an excess of false positives and therefore, the significance of the local parameter estimates may be questionable (A. R. da Silva & Fotheringham, 2016). Therefore, we have also used the newly developed correction method by da Silva and Fotehringham (A. R. da Silva & Fotheringham, 2016) for inference in the GWR/MGWR to solve the multiple testing issues to obtain reliable local parameter estimates. da Silva and Fotheringham proposed an effective correction to the significance level α to assess the significance of local parameter estimates and avoid the proportion of false positives exceeding α. The corrected significance level α value is calculated using **Eq.**(10):

$$\alpha = \frac{\xi}{\frac{ENP}{P}} \quad (10)$$

Where ξ is the expected type I error rate before correction, *ENP* is the effective number of parameters in the model, which is a function of the optimal bandwidth parameter, and *P* is the number of parameters in the model (A. R. da Silva & Fotheringham, 2016).

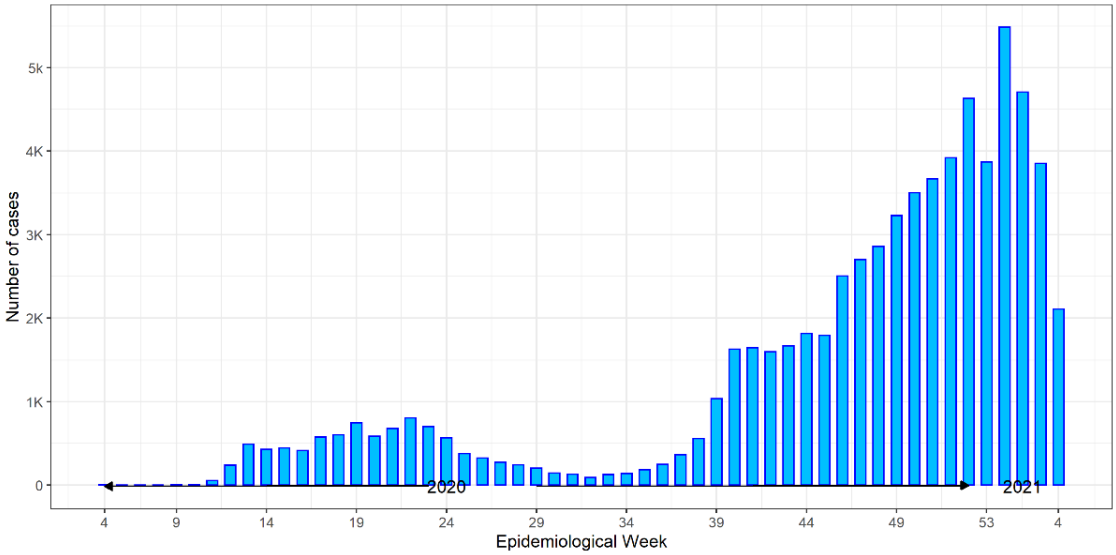
The adaptive bisquare spatial kernel function was applied to develop both local models, and the golden search was applied to select an optimal bandwidth (H. Yu et al., 2020). The adaptive bandwidth is defined as the proportion of data points involved in the calibration process of local estimates and eliminates the influence of outside the neighbourhood spatial units (H. Yu et al., 2020). AICc was used to evaluate and compare the global model fit and performance. The best model fit is indicated by a larger R-square and a smaller AICc value. The outputs from the best-fitted local model were used to map the local parameter estimates, their estimated standard errors, the significant parameter estimates (after adjusting for multiple tests), the local R-square and the local Condition numbers. A condition number greater than 20 can affect model accuracy and inferences, and a value less than 20 indicates no effect of multicollinearity (T. Oshan et al., 2019).

3.4. Results

3.4.1. Descriptive Statistics

Figure 3.2 shows the distribution of the total cumulative sporadic cases of COVID-19 in Toronto by epidemiological week. The total monthly cases started to increase in March (week 12) and experienced a continuous decline from June 2020 (week 25) to August 2020 (week 34). However, the numbers continued to rise exponentially in the following months, with the maximum number of cases witnessed in January 2021 (20,506 total monthly cases).

Figure 3.2. Total sporadic COVID-19 cases by the epidemiological week in Toronto between January 2020 and January 2021.



3.4.2. Spatial and spatiotemporal cluster analyses

3.4.2.1. The Global Spatial Autocorrelation

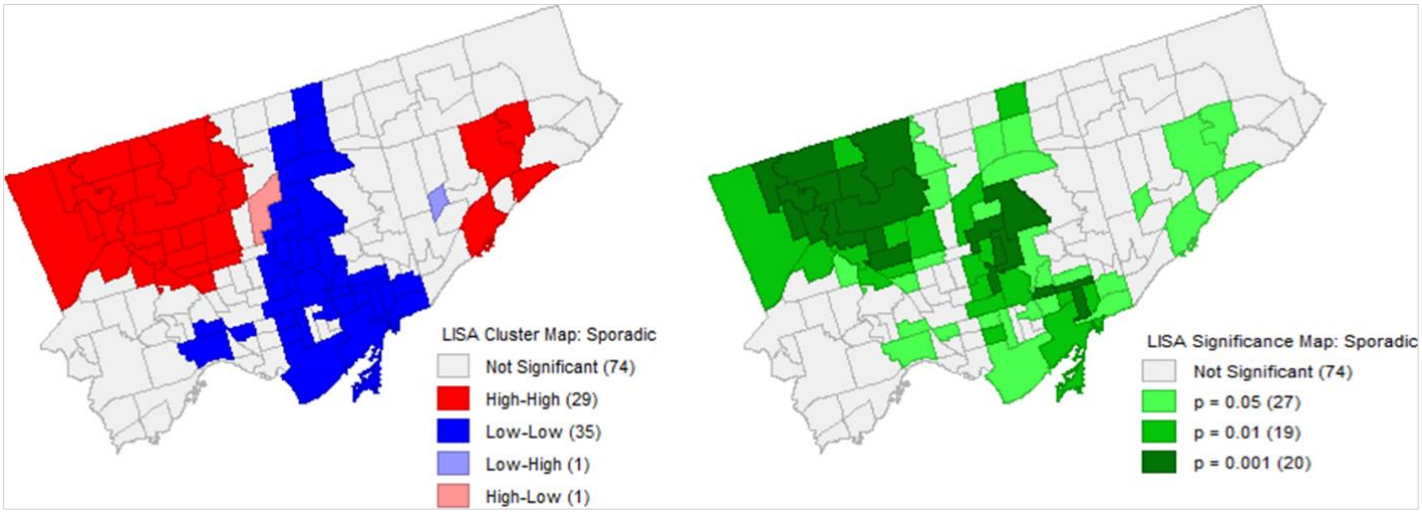
Before implementing the spatial and spatiotemporal cluster analyses, we examined the global Moran's *I* statistic results to evaluate the presence of spatial autocorrelation in the study area (Figure 3.3). The outcome variable (incidence rate of COVID-19) demonstrated a positive spatial autocorrelation suggesting a strong clustering pattern, with a statistically significant Moran's *I* value of 0.62 (p-value=0.001, z-value =12.96). The results indicate that the distribution of the incidence rate of COVID-19 had a positively significant correlation with the incidence rate of the nearest neighbourhoods during the study period.

3.4.2.2. Local Moran's *I*

The results from the local Moran's *I* method show COVID-19 clusters with a significance level of 0.1%, 1% and 5% in the Toronto neighbourhoods. The impositions of the Bonferroni bounds result in only the low-low and high-high clusters being significant. Our goal is to understand the

interesting locations rather than interpreting the most stringent tests(p-value=0.00012), and therefore we have interpreted the use of the traditional p-value of 0.05, 0.01 and 0.001. The results also show distinct clustering of statistically significant high-high clustering or 'hotspots' in the northwestern and southeastern parts of Toronto in 29 neighbourhoods (20.71% of all neighbourhoods) and Low-Low clustering or cold spots in 35 neighbourhoods (25% of all neighbourhoods) in central Toronto (Figure 3.3).

Figure 3.3. Cluster and significance maps of COVID-19 incidence rates in Toronto using the local Moran's *I* approach.



3.4.2.3. The space-time scan statistic

We found eight statistically significant space-time clusters of COVID-19 in different parts of Toronto from January 2020 to January 2021 using the space-time statistic method in SaTScan™ (Figure 3.4, Table 3.1). Table 3.1 provides the space-time cluster characteristics that include the p-values, the total population at risk, the observed and the expected number of cases, the total number of neighbourhoods and the relative risk rates for each cluster. The high-risk clusters varied

in terms of size, the magnitude of relative risks, the total number of neighbourhoods, and the number of populations at risk. Forty-two neighbourhoods (30%) not within the significant clusters were considered low/no risk (of COVID-19) neighbourhoods. The highest relative risks (RR>3.0) were recorded in the western and eastern parts of Toronto. The temporal periods for the eight high-risk clusters fell between October 2020 - January 2021, with cluster 1 (located in the northwestern corner) experiencing a three-month-long clustering period.

Figure 3.4. The space-time clusters of COVID-19 in Toronto between January 2020 and January 2021.

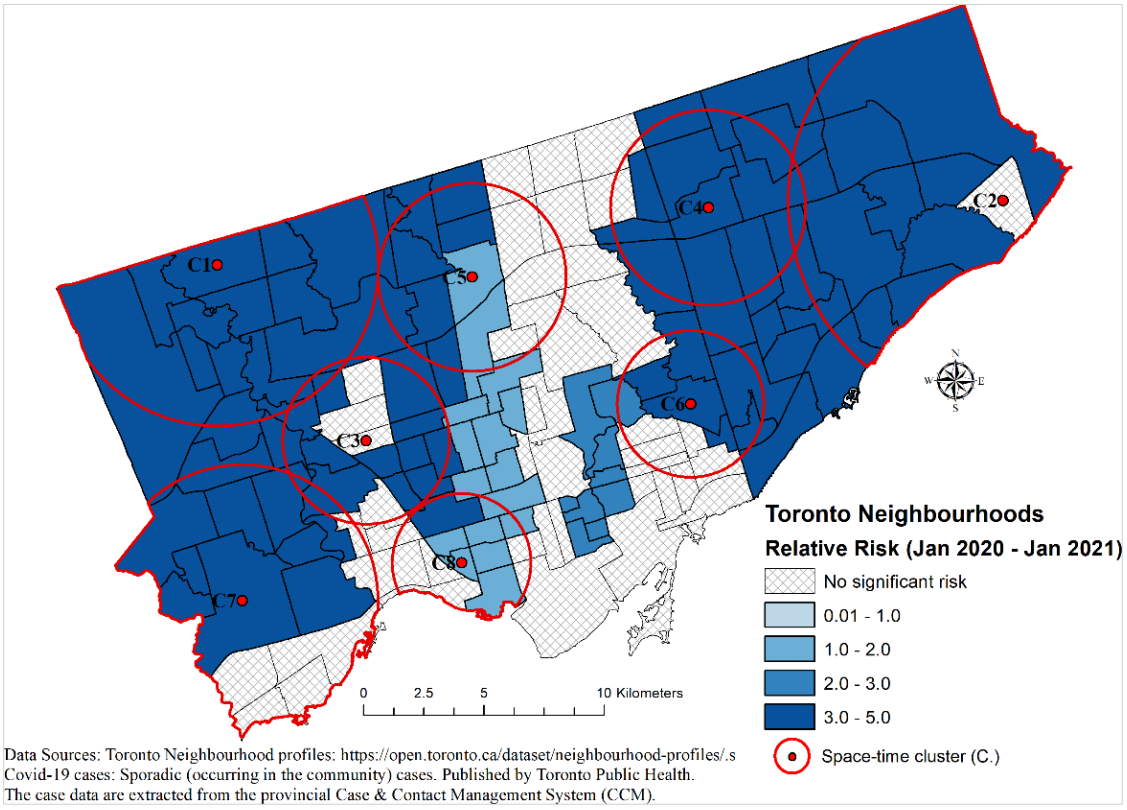


Table 3.1. The characteristics of the COVID-19 space-time clusters in Toronto between January 2020 and January 2021.

CLUSTER	P-VALUE	Radius	START DATE	END DATE	# of Neighbours	OBSERVED	EXPECTED	RR	POP
1	<0.001	6691	2020/10/1	2021/1/31	13	11690	2184.21	6.21	272167
2	<0.001	8912	2020/11/1	2021/1/31	10	7137	1584.07	4.90	263895
3	<0.001	3463	2020/10/1	2021/1/31	16	5967	1765.93	3.60	220046
4	<0.001	4049	2020/11/1	2021/1/31	10	5298	1609.89	3.48	268198
5	<0.001	3905	2020/12/1	2021/1/31	11	3447	987.09	3.62	244011
6	<0.001	3045	2020/11/1	2021/1/31	12	3674	1206.91	3.16	201063
7	<0.001	5643	2020/12/1	2021/1/31	15	2665	1026.03	2.66	253639
8	<0.001	2880	2020/12/1	2021/1/31	10	1608	760.13	2.14	187906

RR = Relative risk, Pop = Population at risk.

3.4.3. Regression Analyses

The results of the bivariate regression analysis show a significant positive correlation between COVID-19 incidences and the four independent variables: percentage of immigrants (p-value<0.0001), the prevalence of low income (p-value<0.0001), unemployment rate (p-value<0.0001) in the neighbourhoods, whereas a significant negative correlation was observed with the percentage of the population with a lower level of education (p-value<0.0001) in the neighbourhood (Table 3.2). The spatial regression models were run with the four independent variables, and the multicollinearity among predictor covariates was tested using the multicollinearity condition number method. The model yielded a multicollinearity condition number of 18.9. Since the multicollinearity condition number was under 30, it indicates that the predictor variables in the regression model are not highly correlated (Anselin et al., 2006; Yandell & Anselin, 1990). The presence of autocorrelation among the predictor variables was

further investigated using a correlation matrix. The results indicated no strong correlation between the four independent variables and were included in the local and global regression models.

Table 3.2. Results of the bivariate regression analysis.

Variables	Estimate (95% CI)	p-value
Percentage of immigrants	**0.69 (0.60, 0.77)**	<0.0001
Percentage of population with lower level of education	**0.82 (0.75, 0.86)**	<0.0001
Prevalence of low income	**0.30 (0.15, 0.45)**	<0.0001
Unemployment rate	**0.64 (0.53, 0.72)**	<0.0001

Note: Outcome variable= log of COVID-19 incidence rate.

3.4.3.1. Global Regression Models

The results from the OLS diagnostics show that the LM lag and LM error were both statistically significant at the 0.05 significance level, indicating that spatial models such as SLM and SEM are more appropriate and relevant global models. The Akaike information criterion (AIC), R-square and the log-likelihood values were compared to find the most appropriate (Yandell & Anselin, 1990) spatial regression model (SLM or SEM) that explains the global spatial relationship between the socioeconomic variables and COVID-19 incidence rates. Table 3.3 shows that the SEM model had the lowest AIC value and the highest R^2value. Therefore, the SEM model is chosen as the most appropriate model that successfully incorporates spatial effects in the model and can better explain the model variability. The results from the SEM model showed a one percent increase of the immigrants in a neighbourhood was associated with a 1% (exponentiation of the estimate of 0.0110) increase in the COVID-19 incidence rate in the neighbourhood. Moreover, a one-unit increase in lower-educated individuals in a neighbourhood was associated with a 2% (exponentiation of the estimate of 0.0179) increase in the COVID-19 incidence rate in the neighbourhood. The spatial autoregressive coefficient (LAMDA) of the SEM model has a positive

sign and is highly significant (p-value<0.05), indicating a positive spatial dependence between the neighbourhoods (Table 3.4).

Table 3.3. Summary of the outputs from the global spatial regression models.

Index	Spatial Lag Model (SLM)	Spatial Error Model (SEM)
R^2	00.79	0.81
Log-likelihood	102.67	107.60
Akaike information criterion	-193.35	-205.21

Table 3.4. Regression outputs from the Spatial Error Model (SEM) model (n=140).

Variable	Coefficient	Std. Error	z-value	Probability
Constant	1.4605	0.1562	9.3487	<0.0001
Percentage of Immigrants	0.0110	0.0035	3.0842	0.0020
Percentage of population with lower level of education	0.0179	0.0026	6.8249	<0.0001
Prevalence of low income	-0.0024	0.0061	-0.3988	0.6899
Unemployment rate	0.0402	0.0255	1.5738	0.1155
LAMDA	0.4455	0.1077	4.1337	<0.0001

Dependent variable: base-10 logarithmically transformed rates of COVID-19
Akaike Information Criterion (AIC): -63.46

3.4.3.2. Local Regression Models (GWR and MGWR)

Table 3.5 shows the covariate-specific bandwidths, the effective number of parameter estimates, critical t values, adjusted alpha, R square and AICc values for both GWR and MGWR models. The diagnostics of the local models indicated that the MGWR model presented the largest R-square and lowest AICc among all models and, therefore, a better-fitted model (Table 3.5). The

R-square value indicates that the MGWR model explains 88.4% of the variations in the COVID-19 rates in Toronto. The summary results of the MGWR are listed in Table 3.6.

Table 3.5. GWR and MGWR Summary Statistics for the COVID-19 Data.

Diagnostic	GWR	MGWR					
	Entire Model	Entire Model	Intercept	Immigrants	Lower level of education	Low Income	Unemployment Rate
Bandwidth	71	n/a	43	65	99	43	139
Effective No. of Parameters	20.23	21.09	6.79	3.83	1.98	7.20	1.28
Adjusted α	0.0125	0.0018	0.007	0.013	0.025	0.006	0.039
Critical t (95%)	2.53		2.72	2.51	2.26	2.74	2.08
AICc	158.105	149.45					
R2	0.874	0.884					

n=140

Table 3.6. Summary of coefficients results from the local MGWR Model.

Variables	**Mean**	**STD**	**Min**	**Median**	**Max**
Intercept	0.019	0.195	-0.248	-0.022	0.526
Percentage of Immigrants	0.307	0.095	0.077	0.310	0.489
Lower level of education	0.505	0.048	0.437	0.492	0.609
Prevalence of low income	-0.019	0.128	-0.368	0.027	0.231
Unemployment rate	0.119	0.008	0.110	0.116	0.136

The spatial associations between the socioeconomic factors and COVID-19 rates using the outputs from the MGWR model are shown in Figures 3.5-3.8, where the local parameter estimates and associated standard errors are mapped in (a) and (b), respectively, while (c) indicates only the significant local parameter estimates defined based on the significance (α) value given in Table 3.4 which has been adjusted for multiple tests.

Significant positive associations between COVID-19 incidence rates and the percentage of immigrants were observed in all neighbourhoods in Toronto (Figure 3.5c). The higher local parameter estimates were observed in the northwestern and southern parts of Toronto, whereas lower local parameter estimates were observed in the northeastern parts of Toronto. Similar to the percentage of immigrants, a higher percentage of the population with a lower level of education was positively associated with COVID-19 incidence rates throughout all neighbourhoods of Toronto (Figure 3.6c). The higher local parameter estimates were observed in northeastern Toronto, whereas relatively lower local parameter estimates were observed in southern Toronto.

The association between COVID-19 incidence rate and the two factors: prevalence of low income and unemployment rate, were not found to be significant in the majority of the neighbourhoods in Toronto (Figures 3.7c and 3.8c). The prevalence of low-income local parameter estimates varied from positive to negative values (Figure 3.7a). However, only six neighbourhoods in northwestern Toronto had a significant negative association with COVID-19 incidence rates (Figure 3.7c). The unemployment rate was found to be significantly associated with COVID-19 incidence rates in only five neighbourhoods in the northeastern part of Toronto (Figure 3.8c)

Figure 3.9 presents spatial variations in local R-square values in the study area and the local condition numbers (CN) from the MGWR model for each neighbourhood. The highest local R-square values ($R^2 \geq 0.86$) were observed in the western and central parts of Toronto, namely in the Etobicoke and North York regions (Figure 3.9a). The local condition number (CN) was observed higher in the northern and central parts of Toronto. However, the overall the CN

value was <20 with maximum value being 9.5, suggests that there were no presence of multicollinearity in the model (Figure 3.9b).

Figure 3.5. (a) Local parameter estimates (b) Standard errors (c) Significant local parameter estimates (after adjusting for multiple tests) for the percentage of immigrants variable from the MGWR model.

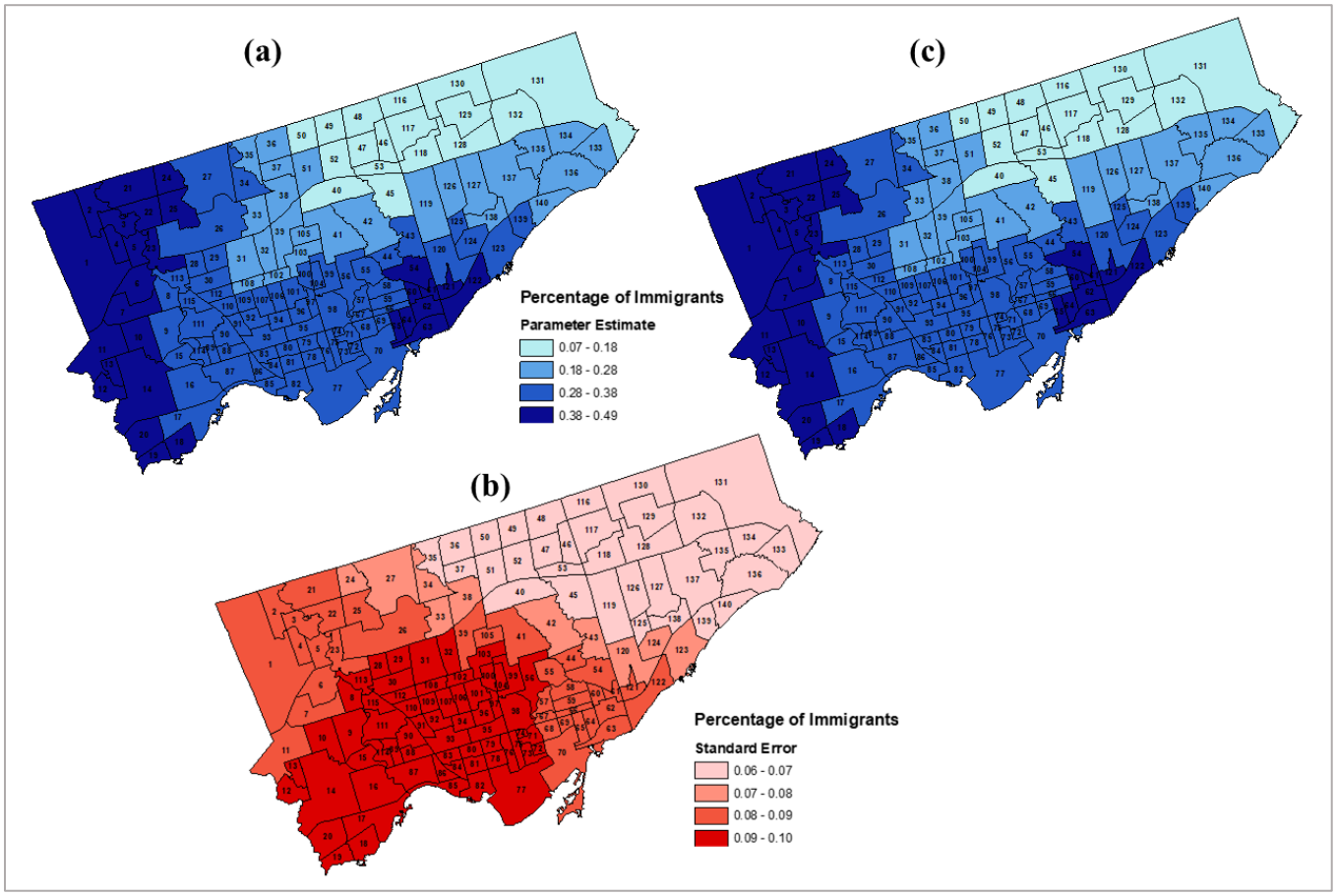


Figure 3.6. (a) Local parameter estimates (b) Standard errors (c) Significant local parameter estimates (after adjusting for multiple tests) for the lower level of education variable from the MGWR model.

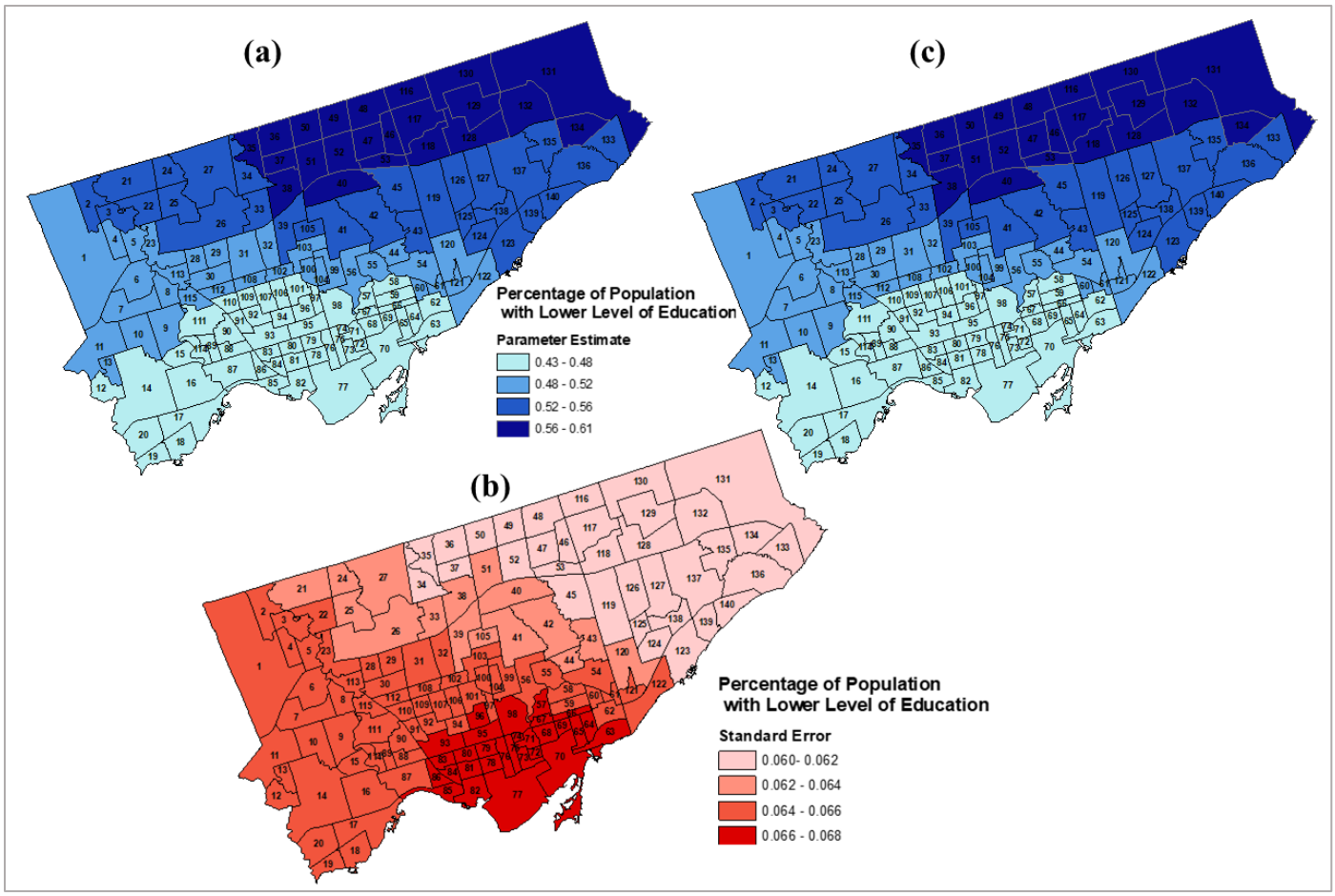


Figure 3.7. (a) Local parameter estimates (b) Standard errors (c) Significant local parameter estimates (after adjusting for multiple tests) for the prevalence of low-income variables from the MGWR model.

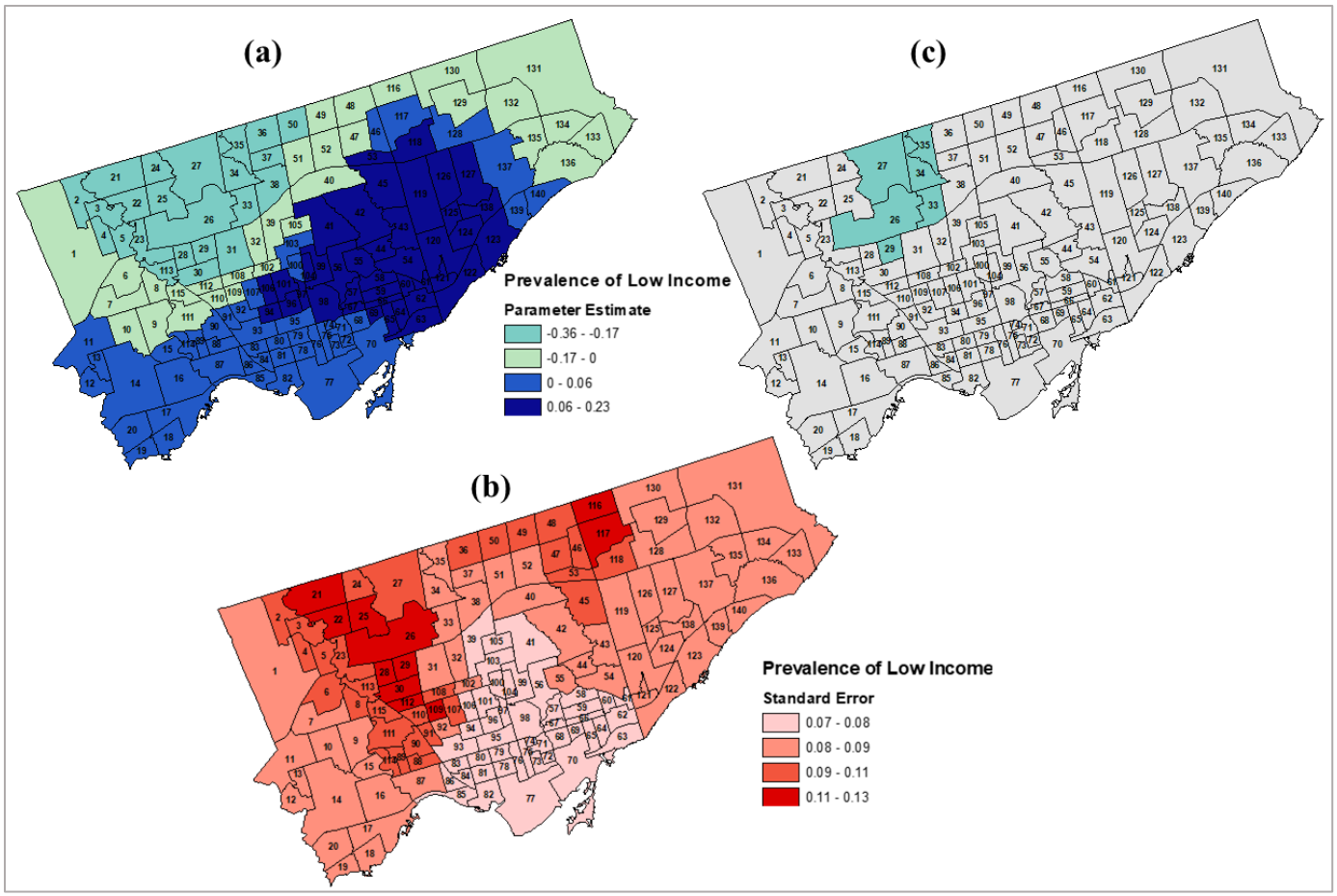
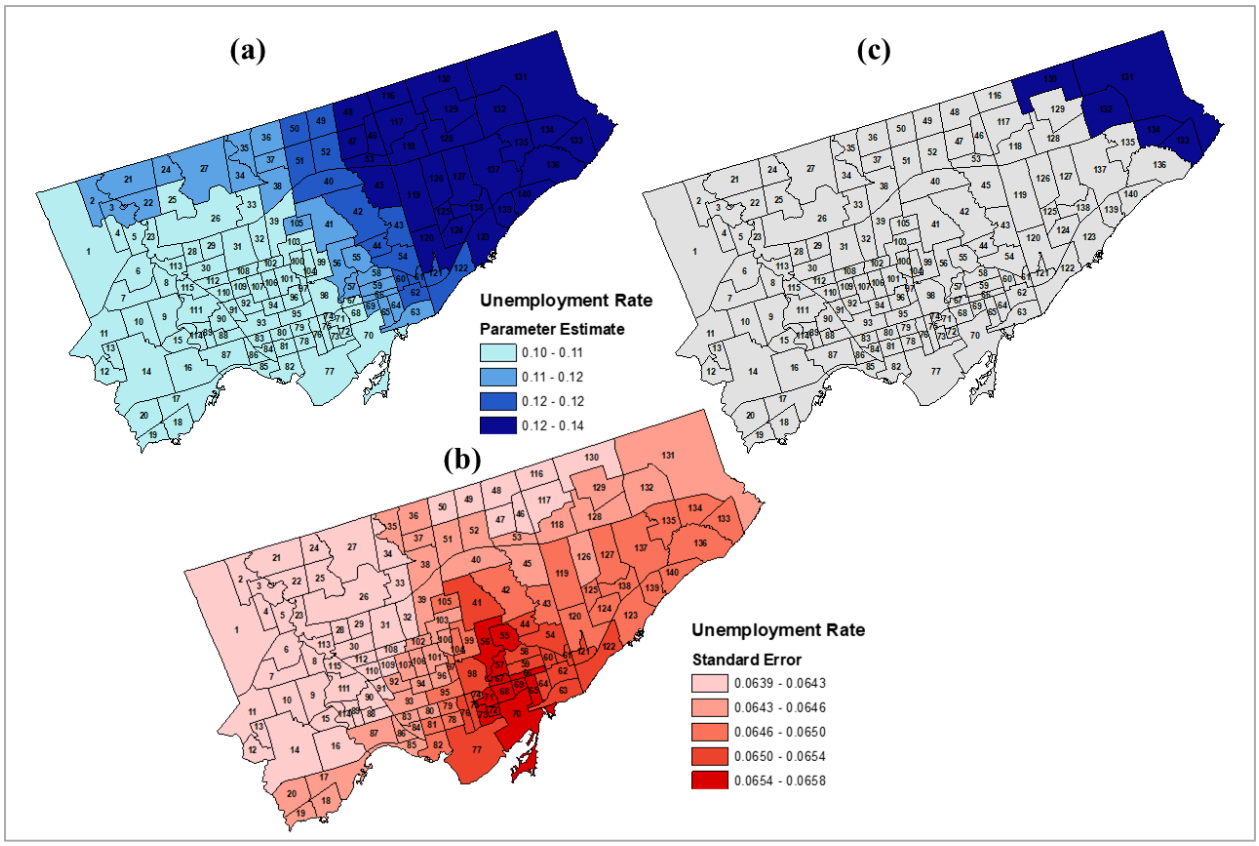


Figure 3.8. (a) Local parameter estimates (b) Standard errors (c) Significant local parameter estimates (after adjusting for multiple tests) for the unemployment rate variable from the MGWR model.



3.5. Discussion

The results of our analysis of the COVID-19 incidences displayed non-random spatial distribution patterns in our study area in Toronto. Both Anselin's local Moran's *I* and Kulldorff's space-time statistic produced similar patterns of hotspots/clusters in the Northwestern part of Toronto in the Etobicoke region. An assessment of the MGWR maps suggests that the higher percentage of immigrants in the neighbourhoods may have contributed to the higher COVID-19 incidence in that part of Toronto. Additionally, in the eastern part of Toronto, the lower level of

education of the people in the neighbourhood may have contributed to higher COVID-19 incidence.

Although there were some similarities, we also observed a few dissimilarities in the clusters detected by the two clustering approaches. It is important to note that the local Moran's *I* do not take the temporal factors into account, and the magnitude of the risk is not provided, which is covered in the scan statistic. While the scan statistic identified clusters in the southwestern parts of Toronto, the local Moran's *I* found no significant clusters in those areas. Overall, the clusters identified by the scan statistic were more localized and covered a broader area than that identified by the local Moran's *I*. In local Moran's *I* method, the clusters are identified in a strictly bounded area, where the correlation is assessed between the disease rate of a certain neighbourhood and the average disease rate of its surrounding (first-order contiguity) neighbourhoods (Laohasiriwong et al., 2018). Scan statistic might be a more sensitive method (Laohasiriwong et al., 2018) and provides more elaborate details of the cluster characteristics such as the total population at risk, radius, and relative risks compared to local Moran's *I*. Since each method has its own set of strengths and limitations, we believe a combined approach to identifying the disease clusters would provide an in-depth understanding of the disease clusters and ensure logical consistency of the results of the analyses.

The results of the global regression yielded that the percentage of immigrants and a lower level of educated people in the neighbourhood influenced to vary COVID-19 incidence rates in Toronto, consistent with the previous studies (Borjas, 2020; Fielding-Miller et al., 2020; Goutte et al., 2020). We found that a higher concentration of immigrants in the neighbourhood was associated with the increased COVID-19 incidence in the neighbourhood, suggesting that

COVID-19 might have disproportionately impacted the immigrant population in Toronto. This could be attributed to immigrants living in overcrowded housing conditions, lower-income neighbourhoods and/or working in an environment where physical distancing is often challenging (Choi et al., 2020; CTV, 2020). It is worth noting that Asian countries account for the majority (53%) of immigrants in Toronto, with China alone accounting for 20%. Heterogeneity in risk factors across Asian-origin subgroups may have led to distinct patterns of geographic disparities of risk in COVID-19.

We also found that neighbourhoods with lower levels of education had higher outbreaks of COVID-19 in Toronto. This suggests that the least educated people may have a lower understanding and awareness of COVID-19, highlighting the importance of knowledge to fight this pandemic. It is also possible that the less educated people are employed in professions where they are at higher risk of contracting COVID-19, and it is not possible for them to self-isolate at home. However, note that the global models have some limitations since these models assume that the interactions between the variables are stationary over space which is often not realistic (Thayn & Simanis, 2013). A variable that is not found significant at the global scale may not be true at a local spatial scale. The results of our study also suggest that although the prevalence of low income and the unemployment rate were not found to be significant predictors in the global model, the individual influences of these two predictors exhibited a significant association (negative and positive) in the local MGWR model in a few neighbourhoods in the city. These suggest that ignoring the locally variable indicators may not help to control disease transmission.

A major advantage of using the local regression model in this study is that it allows us to visually demonstrate the magnitude of risk deriving from a factor at a spatial scale, which would

allow us to know where a particular type of intervention will be required based on the different associations between characteristics of given populations located at different places. The intervention may also include informing on COVID protection measures using which languages and where to ask people to stay at home and work remotely, etc. The parameter estimates and significant maps allowed a comparison with the cluster maps to identify the socioeconomic drivers of the clustering of the disease in neighbourhoods.

While the global models performed well, the local models provided a more parsimonious quantitative output of the socioeconomic determinants that may influence COVID-19 rates. The MGWR allowed the relationship between COVID-19 and explanatory variables to vary spatially and at different scales. The MGWR has the advantage of more accurately depicting spatial heterogeneity, diminishing multicollinearity and lessening the bias in the parameter estimates (T. M. Oshan et al., 2019; H. Yu et al., 2020). Our findings indicated that the MGWR does not suffer from the effect of multicollinearity and is robust as it allowed each parameter to be processed at flexible and varying scales. The use of the correction method proposed by da Silva and Fotheringham (A. R. da Silva & Fotheringham, 2016) allowed us to identify the significant association with each predictor variable at a local scale for a more validated output. The MGWR model explained 88.4% of the model variances, which is higher than the global four-parameter SEM model (81%), SLM (79%), and the local GWR model (87.4%), suggesting that the results of the MGWR model provided a more enhanced model accuracy. Therefore, we believe using the MGWR approach to demonstrate the influence of socioeconomic components on COVID-19 incidence patterns in Toronto benefitted our study.

There are several limitations to our study. First, we used a limited number of variables based on the publicly available data at the neighbourhood level. Further study by including additional explanatory variables may benefit our understanding of the spatial variations of the disease incidences. Second, this study used confirmed and probable COVID-19 cases collected from Toronto Public Health. However, COVID-19 is often known to be asymptomatic. Thus, we could have missed the infected individuals who had mild symptoms and not visited any hospital or testing centers and were unreported.

Furthermore, there is a possibility of misclassification and variation in the propensity to test. However, we believe that these unreported cases could be randomly distributed and may not have a significant impact on our analysis. Third, the scan statistic uses a circular window to detect disease clusters and is unable to detect clusters irregular in shape (Hughes & Gorton, 2013; Kulldorff & Information Management Services Inc, 2009).

In Toronto, COVID-19 has rapidly evolved, creating a dire public health crisis. Several ongoing preventative measures were adopted in Toronto by the government to control the spread of COVID-19 due to the higher rates of infections among the residents. However, the number of cases continued to rise. Spatial modelling of disease is important to assess a newly emerging infectious disease such as COVID-19 to understand the magnitude of the risk in a densely populated city like Toronto. Some studies have emphasized the importance of early implementation of intervention strategies to mitigate the disease risk from COVID-19 (Pei et al., 2020; Pellis et al., 2020). Our study has several policy implications for mitigating the risk from COVID-19: i) it may serve as a spatial guideline for the decision-makers to formulate mitigation strategies focusing on the hotspots across the neighbourhoods with effective and clear guidelines,

ii) the socioeconomic determinants of COVID-19 may provide spatially explicit information about the spatial drivers of COVID-19 to identify the localities, and the policymakers can establish disease surveillance based on the socioeconomic drivers that were influencing disease risk in the neighbourhoods, iii) it may help in developing plans for decreasing the socioeconomic inequalities in the high-risk neighbourhoods to mitigate the disease risk, and iv) it may help to identify areas where interventions will be required to improve public knowledge and awareness of COVID-19.

Future research may focus on how the epidemics are disseminated by using diffusion modelling techniques to display the diffusion direction, magnitude, and dynamics by taking into account of the knowledge related to the virus, such as the incubation time and injected into modelling to better represent the reality.

3.6. Conclusion

The use of spatial methods to model COVID-19 incidence in Toronto is warranted to improve the current control and vaccination strategies at the postal code level. Our study adopts multiple spatial and spatiotemporal models to provide a deeper insight into the magnitude of the risk of COVID-19 in Toronto. We detected several high-risk clusters in different parts of Toronto, and the socioeconomic conditions in the neighbourhoods could be the underlying factors for clustering the cases. Addressing the socioeconomic determinants of health, such as education, diversity, income, and unemployment status, is important in infectious disease surveillance. People living in a lower socioeconomic condition may struggle to adhere to proper social distancing measures. The findings of our study could allow a closer focus on the COVID-19

incidence and the socioeconomic predictors to mitigate the disease risk and control it. The policymakers could be beneficial from the findings of this study.

Chapter 4 Manuscript 2: Identifying Spatiotemporal Patterns of COVID-19 and the Drivers of the Patterns in Toronto: A Bayesian Hierarchical Spatiotemporal Modelling[2]

4.1. Summary

Spatiotemporal patterns and trends of COVID-19 at a local spatial scale using Bayesian approaches are hardly observed in the literature. Also, a study rarely used satellite-derived long time-series data on the environment to predict COVID-19 risk at a spatial scale. In this study, we modelled the COVID-19 pandemic risk using a Bayesian hierarchical spatiotemporal model that incorporates satellite-derived remote sensing data on land surface temperature (LST) from January 2020 to October 2021 (89 weeks) and several socioeconomic covariates of the 140 neighbourhoods in Toronto. The spatial patterns of risk were heterogeneous in space with multiple high-risk neighbourhoods in Western and Southern Toronto. Higher risk was observed during Spring 2021. The spatiotemporal risk patterns identified 60% of neighbourhoods had a stable, 37% had an increasing, and 2% had a decreasing trend over the study period. LST was positively, and higher education was negatively associated with the COVID-19 incidence. We believe the use of Bayesian spatial modelling and the remote sensing technologies in this study provided a strong versatility and strengthened our analysis in identifying the spatial risk of COVID-19. The findings would help in the prevention plan, and the framework of this study may be replicated in other highly transmissible infectious diseases.

[2] The citation for this article is Nazia, N., Law, J. & Butt, Z.A. Identifying spatiotemporal patterns of COVID-19 transmissions and the drivers of the patterns in Toronto: a Bayesian hierarchical spatiotemporal modelling. *Sci Rep* **12,** 9369 (2022). https://doi.org/10.1038/s41598-022-13403-x

4.2. Introduction

COVID-19, caused by the coronavirus SARS-CoV-2, has complex transmission dynamics possibly generated by different risk factors such as demographic, social and environmental factors (Cao et al., 2020; Franch-Pardo et al., 2020; Guan et al., 2020; P. Wang et al., 2021). It is highly transmissible by either direct contact with an infected individual or transmission via contaminated surfaces leaving the world at a halt in many aspects (Ganegoda et al., 2021). In Canada, over 3.4 million COVID-19 cases and over 37,485 COVID-19-related deaths have been reported, with Ontario and Quebec reporting the highest cumulative cases in the nation (Government of Canada, 2020). Over 11.2 billion COVID-19 vaccine doses have been administered around the world (WHO, 2021), and at least 89% population over the age of 5 years in Canada have received at least one dose (March 28, 2022) (Canada, 2021a). In Canada, Toronto continues to experience substantial COVID-19 incidence and hospitalization rates despite several interventions and mitigation efforts made by local and provincial public health officials (CBC News, 2021; CTV News, 2021).

As momentum grows to end this global pandemic, understanding the disease trends, detecting hotspots, and identifying important risk factors at the community level is an imperative research effort. Temperature is often a significant risk factor for infectious diseases because a certain temperature can help the virus evolve rapidly (Tobías & Molina, 2020). A county-level study in the USA found a strong negative influence of nighttime land surface temperature (LST) with COVID-19 using low-resolution Moderate Resolution Imaging Spectroradiometer (MODIS) images of 2020 (Johnson et al., 2021). Another study found LST to be an important determining factor in the COVID-19 infection rate in Kolkata, India (Das et al., 2020). Another

study by Hassan et al. identified a strong positive relationship between COVID-19 and LST (Hassan et al., 2021). This study showed that a 1°C increase in LST is linked with a 36.1% increase in COVID-19 incidence rates in Bangladesh. Many prior studies have discussed the associations between temperature and COVID-19 (Shi et al., 2020); however, most of these studies have used short temporal periods, lower resolution images or have not performed a small area analysis. The results were diverse and often contradictory in different geographic areas, leaving a gap in understanding the impact of temperature on COVID-19 transmission in a small urban area.

Previous studies have also linked different socioeconomic and demographic factors to explain the heterogeneity in COVID-19 rates across space. Some studies have found that areas with low socioeconomic status, such as rate of poverty (Abedi et al., 2020; Chaudhry et al., 2020; Cordes & Castro, 2020; Maiti et al., 2020; Sannigrahi, Pilla, Basu, Basu, et al., 2020b), rate of education (Abedi et al., 2020; Cordes & Castro, 2020; Goutte et al., 2020; Y. Wu et al., 2020), ethnicity status (Andersen et al., 2021b; Y. Chen & Jiao, 2020b; Cordes & Castro, 2020; Kathe & Wani, 2020; Maiti et al., 2020; Niedzwiedz et al., 2020; F. Sun et al., 2020b; Y. Sun et al., 2020b), immigration status (Borjas, 2020; Fielding-Miller et al., 2020; Indseth et al., 2021; Ingen et al., 2021; Politi et al., 2021), unemployment rate (Goutte et al., 2020; Y. Sun et al., 2020b; Yoshikawa & Kawachi, 2021), and housing conditions (Borjas, 2020; Goutte et al., 2020; Y. Wu et al., 2020) tend to experience higher rates of COVID-19 infections and morbidity due to the economic and health inequalities. A previous work by Vaz in 2021 (Vaz, 2021) found that COVID-19 cases are not uniformly distributed in Toronto. Social injustice, socioeconomic vulnerability and population density were found to be related to the increasing spread and

incidence of COVID-19. A work by Feng in 2021 has found that neighbourhood-level population density and low income have a significant effect on COVID-19 mortality risk (C. Feng, 2021). Another study of Toronto neighbourhoods by Choi et al. in 2021 found that several demographic and socioeconomic factors, such as higher education rate, and lower rates of immigrants (foreign-born residents) were significantly associated with decreasing the number of COVID-19 infections (Choi et al., 2021). These past studies implied that these factors might disproportionately impact COVID-19 infection rates. Finally, analyzing the spatiotemporal trends of COVID-19 transmission to understand whether the disease risk trends show increasing, decreasing or stable patterns over the study period has also been understudied.

In this study, we used Bayesian hierarchical spatiotemporal models to investigate the spatiotemporal patterns of COVID-19 transmission in Toronto. The approach allows us to deal with uncertainties related to the data, the process and model parameters (Haining & Li, 2020) since it has the capacity to account for missing data, measurement errors and ecological bias (Knorr-Held & Raßer, 2000; Law et al., 2014). Even though Bayesian models have a clear advantage, only a handful of studies (Jaya & Folmer, 2021; MacNab, 2022; Rohleder & Bozorgmehr, 2021; Saavedra et al., 2021; L. Wang et al., 2021) in COVID-19 research have adopted Bayesian approaches to predict COVID-19 risk, identify trends and locate hotspots. Motivated by the recent increase in COVID-19 incidence in winter, this study scrutinizes the effect of weather and socioeconomic and demographic factors on COVID-19 using a small area analysis. Within the scope of this study, we will answer four research questions: i) where were the hotspots of COVID-19, ii) what were the temporal patterns of risk in the study area, iii) was there a relationship between land surface temperature and COVID-19 while adjusting for socioeconomic and demographic

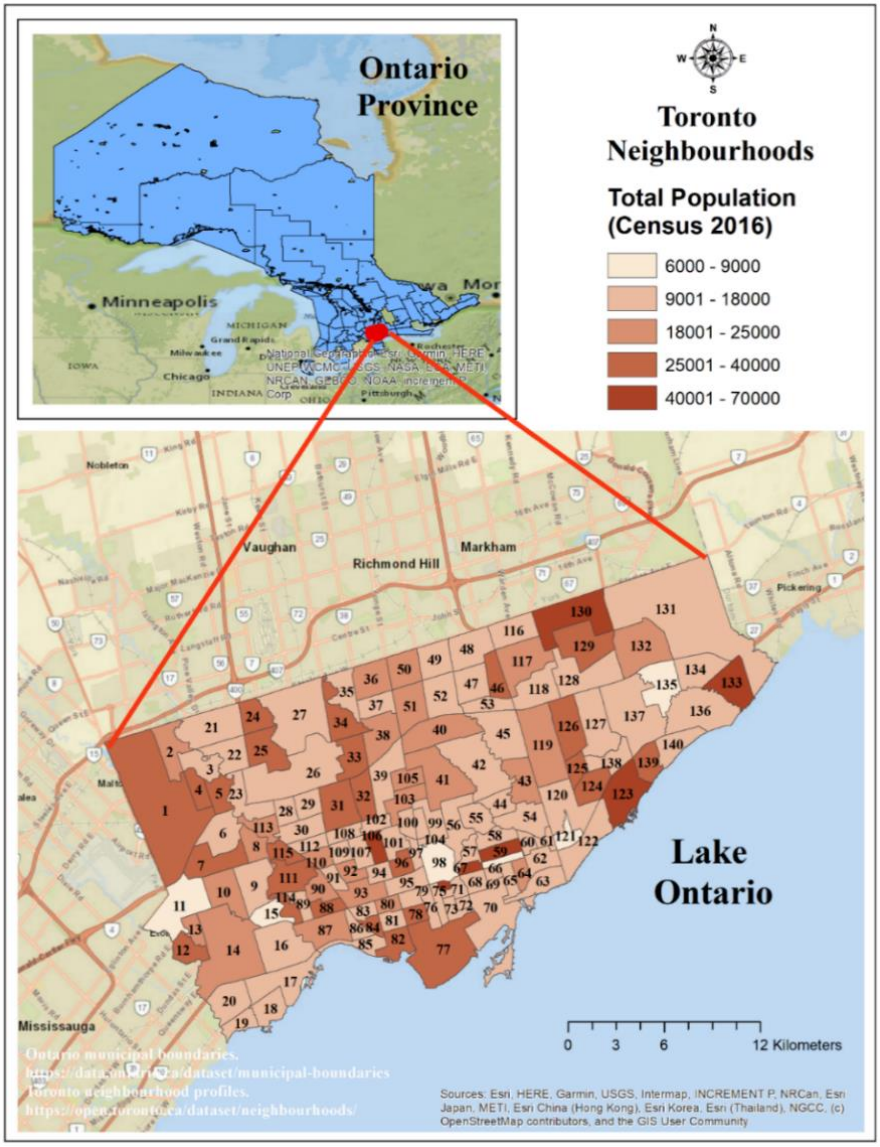
factors, and iv) was there any spatiotemporal trend of COVID-19 transmission in Toronto (e.g., stable, increasing or decreasing)?

4.3. Methods

4.3.1. The Study Area

The study area is the metropolitan city of Toronto, located on the northwestern shore of Lake Ontario at an altitude of 175m (43°42′00″ N latitude and 79°24′58″ W longitude). It is the capital city of the province of Ontario in Canada, with a total land area of 630 km^2 and a high population density of 4,692 persons/km^2. Toronto has a well-defined urban heat island with warmer temperatures, mostly at night and in winter, compared to the rest of the city's surrounding regions. The average temperature of Toronto is 21.9°*C* (81.3°F), and the annual rainfall is 845mm (33.3 inches), with July *(17°C to 25°C/62°F to 77°F)* the hottest and February (average -4.4 °C /24.1 °F) the coldest months of the year. The map of the study area is provided in Figure 4.1.

Figure 4.1. The study area in Toronto, Ontario, Canada.



Note: The numbers inside the neighbourhoods represent the neighbourhood ID.

4.3.2. Geographic boundary, population, and case data

There were 140 geographically distinct neighbourhoods in the study area. The digital data of the geographic boundaries for these neighbourhoods were acquired from the open data portal of the city of Toronto (City of Toronto, 2020b). We used the COVID-19 case data originally collected

by Toronto Public Health and extracted from the provincial Case & Contact Management System (CCM) by the city of Toronto. The case dataset contains the demographic, geographic, and severity information for all confirmed and probable, sporadic, and outbreak-associated cases. We aggregated the daily case data from January 2020 to October 2, 2020, by the 89 epidemiological weeks and at the neighbourhood level. The total population for each neighbourhood was collected from neighbourhood profiles using the census 2016 population dataset collected and released by Statistics Canada (City of Toronto, 2017).

4.3.3. Demographic and socioeconomic variables

Previous studies have also linked different socioeconomic and demographic factors to explain the heterogeneity in COVID-19 rates across space (Abedi et al., 2020; Andersen et al., 2021b; Chaudhry et al., 2020; Y. Chen & Jiao, 2020b; Cordes & Castro, 2020; Das et al., 2020; Fielding-Miller et al., 2020; Goutte et al., 2020; Hassan et al., 2021; Kathe & Wani, 2020; Maiti et al., 2020; Niedzwiedz et al., 2020; Sannigrahi, Pilla, Basu, Basu, et al., 2020b; F. Sun et al., 2020b; Y. Sun et al., 2020b; Y. Wu et al., 2020). Based on these past studies and data availability, we selected 11 potential demographic and socioeconomic variables under six categories (demographic, core housing need, race/ethnicity/minority status, education, economic and immigration status) for our model. These data were collected from the Toronto neighbourhood profile based on the 2016 census from Statistics Canada. The full descriptions of the variables are provided in Table 4.1.

Table 4.1. Descriptions of the demographic and socioeconomic covariates (2016 Census).

Category	Covariate	Covariate Description
Education	Higher education rate	Percentage of the population who are aged 25-64 years and have a higher level of education (having a university certificate, diploma or at least a bachelor's degree) (Government of Canada, 2017b).
Economic	Unemployment rate	Percentage of population in private households who are over 15 years and not employed in the labour force) (Government of Canada, 2017f).
	Prevalence of low-income	The percentage of the population whose income falls below the low-income cut-off (LICO) table represents the poverty line (Government of Canada, 2017d, 2017e). The cut-off thresholds represent the income levels at which these families or persons were expected to spend 20 percentage points or more of their after-tax income than average on food, shelter and clothing (Government of Canada, 2017d).
Core Housing Need	Rate of unaffordable housing	Percentage of households in a neighbourhood costs ≥30% of total before-tax income ton adequate shelter (Government of Canada, 2017h).
	Rate of inadequate housing	Percentage of population in the neighbourhood not requiring any major repairs as reported by the residents (Government of Canada, 2017h).
	Rate of unsuitable housing	Percentage of population in the neighbourhood without suitable accommodations according to the National Occupancy Standard (NOS) standards (Government of Canada, 2017h).
Race/Ethnicity/Minority Status	Percentage of black population	The percentage of black population in a neighbourhood.
	Visible minority	Percentage of visible minority in a neighbourhood. Visible minority refers to a person who belongs to a visible minority group as defined by the Employment Equity Act, defining visible minorities as "persons, other than Aboriginal peoples, who are non-Caucasian in race or non-white in colour (Government of Canada, 2017g).
	Immigrants	Percentage of the population born outside of Canada and who is, or who has ever been, a landed immigrant or a permanent resident (Government of Canada, 2017c).
Demographic	Population density	The number of persons per square kilometre in a neighbourhood.

4.3.4. The Land Surface Temperature (LST) Data

We used Landsat imageries to retrieve the weekly average land surface temperature (LST) at the neighbourhood level for 89 consecutive weeks (January 19, 2021 – October 2, 2022). The timing of these imageries was approximately 4 PM (GMT). We collected mostly Landsat 8 imageries (75), and if the imageries for a particular week were not available in Landsat 8, we used Landsat 7 imageries instead. However, no images were available for two epidemiological weeks (week 45 of 2000 and week 7 of 2021). We used the average temperature of the previous and following weeks for these two weeks. The images were corrected using atmospheric correction parameters collected by the National Aeronautics and Space Administration (NASA) to improve estimation accuracy. The complete details of the LST retrieval process, image data, acquisition time, and atmospheric parameters are summarised in Table S4.1 in Appendix 1 (Supplementary Materials).

4.3.5. Zonal Statistics

Zonal statistics were performed in ArcGIS Desktop software to calculate the average LST values for the retrieval of spectral radiance. We applied a mask comprised of the polygons (neighbourhood boundaries) from the map of the city of Toronto and used zonal statistics to calculate the average LST of all pixels by neighbourhood.

4.3.6. Variable Selection

Variable selection was conducted using a two-step method to fit into the multivariable regression model. In the first step, we performed a Pearson's correlation and generated a correlation matrix taking into account of all the potential risk factors to evaluate collinearity among these risk factors (Appendix 2, Figure S4.1 in Supplementary materials). Note that the presence of collinearity among the independent variables can result in model overfitting, unstable estimates and inaccurate

variances, and consequently incorrect inferences about associations between explanatory and the response variables.(Mason & Perreault, 1991; Midi et al., 2010; D. C. Wheeler, 2007; D. Wheeler & Tiefelsdorf, 2005) We observed that the percentage of immigrants had a high correlation (>0.7) with the percentage of the black population and the percentage of visible minorities. Since Choi et al. in 2021, in a study conducted in our study area in Toronto, stated that the percentage of immigrants is an important risk factor in our study area in Toronto (Choi et al., 2021), we selected this variable over the percentage of the black population and percentage of visible minorities. The prevalence of low income was found to be strongly correlated with the rate of unaffordable housing and unemployment rate. Out of the three core housing need variables, the rate of unsuitable and the rate of unaffordable housing had a strong correlation with multiple variables. Since the rate of inadequate housing did not have a strong correlation with any other variables, we selected this (inadequate housing) variable over other housing variables. Among the correlated factors, we selected the prevalence of low income over the unemployment rate based on an earlier study in our study area that found low income to be strongly associated with COVID-19 (C. Feng, 2021). Finally, since the LST and population density did not have a strong correlation with other factors, we kept both in the model.

In the second step, the six selected variables from the first step: Land Surface Temperature (LST), the prevalence of low income, rate of higher education, inadequate housing, percentage of immigrants and population density per square kilometres, were fitted in a Bayesian variable selection method using BayesVarSel package in RStudio to select the variables that fit best in our Bayesian hierarchical spatial model (Appendix 2 in Supplementary Materials). The approach uses priors as proposed by Bayarri et al. (Bayarri et al., 2012), computes posterior probabilities of

hypotheses or the models, and delivers tools in a coherent and complete analysis to properly summarise the outputs (Garcia-Donato & Forte, 2018). This approach yielded LST, rate of higher education and percentage of immigrant variables with higher posterior probabilities and marginal inclusion probabilities (Appendix 2, Table S4.2), suggesting that these three variables are very relevant, highly influential, and the best fit for our Bayesian regression models.

4.3.7. Standardization of the variables

Since the variables were in different units, such as raw values, percentages, and prevalence rates, we used the Z-transformation technique, where the mean for all values was subtracted from each value and was then divided by the standard deviation of the values of the variables to obtain standardized values for the Bayesian model.

4.3.8. Bayesian Spatiotemporal Models

We performed four Bayesian hierarchical space-time models to investigate the long-term spatiotemporal effects of COVID-19 using two frameworks: space-time separable (Haining & Li, 2020) and the space-time inseparable models.(Haining & Li, 2020; Knorr-Held, 2000) Model 1 drew the space-time separable framework, while Models 2, 3 and 4 drew the space-time inseparable modelling frameworks.

To model each outcome value y_{it}, the COVID-19 case count observed in week *t* in the neighbourhood *i* (*i* =1,….*N* and *t* =1,…..*T*), the data model takes Poisson distributions as the likelihood in Eq. 1:

$$y_{it} \sim Poisson\ (\mu_{it}) \quad (1)$$

Specifically, Poisson mean, μ_{it} is a product of n_i, the total number of populations in neighbourhood *i, o*btained from the 2016 census, is assumed to be time-variant, and θ_{it}, the underlying unknown COVID-19 risk in the neighbourhood *i* during week *t*. The space-time

variability is partitioned into three components: spatial, temporal and the space-time interaction effect. In Model 1, a space-time separable model is used that consists of the first two components (spatial, and temporal), where the variability of data is not captured by the space-time separable structure. Models 2,3 and 4 capture the space-time inseparable structure proposed by Knorr-held (Knorr-Held, 2000), using three different space-time interaction effects that allow space-time inseparability. Table 4.2. and Appendix 3 (Supplementary materials) summarise the four Bayesian Space-Time Hierarchical models, including data, process, space-time interaction components, and full model specifications.

Table 4.2. Summary of the four Bayesian Space-Time Hierarchical models.

	Model 1	**Model 2** Type I	**Model 3** Type II	**Model 4** Type III
Framework	Space-Time Separable	Space-Time Inseparable		
Data Model	$y_{it} \sim Poisson\ (\mu_{it}), \mu_{it} = n_i . \theta_{it}$ where Poisson mean, μ_{it} is a product of n_i, the total number of populations in neighbourhood *i* is assumed to be time-variant, and θ_{it}, the underlying unknown COVID-19 rate in neighbourhood *i* during week *t*.			
Process Model	$\log(\theta_{it}) = \propto + beta1 * EDU_{it} + beta2 * IMMI_{it} + beta3 * LST_{it} + (S_i + U_i) + v_t$ (1)	$\log(\theta_{it}) = \alpha + beta1 * EDU_{it} + beta2 * IMMI_{it} + beta3 * LST_{it} + (S_i + U_i) + v_t + \delta_{it}$		
The Overall Spatial Components	$S_{1:140} \sim ICAR(W_{sp}, \sigma_S^2)$ $U_i \sim N(0, \sigma_U^2)$			
Overall Temporal Component	$v_{1:89} \sim ICAR(W_{RW1}, \sigma_v^2)$			
The space-Time Interaction component	$\delta_{it} = 0$	$\delta_{it} \sim N(0, \sigma_\delta^2)$	$\delta_{i,1:T} \sim RW(W_{RW1}, \sigma_\delta^2)$	$\delta_{i:N,t} \sim ICAR(W_{SP}, \sigma_\delta^2)$
The Parameter Model	$\sigma_s \sim Uniform(0.0001,10), \sigma_u \sim Uniform(0.0001,10),$ $\sigma_v \sim Uniform(0.0001,10), \sigma_\delta \sim Uniform(0.0001,10),$ $\propto \sim Uniform(-\infty, +\infty)$			

Other Vague Priors	$beta1 \sim Normal(0,0.0001)$, $beta2 \sim Normal(0,0.0001)$, $beta3 \sim Normal(0,0.0001)$

Where $\boldsymbol{beta1}$, $\boldsymbol{beta2\ and\ beta3}$ are the regression coefficients of higher education rate, percentage of immigrants and LST, respectively.
$\boldsymbol{S_i}$ is the structured, $\boldsymbol{U_i}$ is the unstructured, $\boldsymbol{v_t}$ is the temporal and $\boldsymbol{\delta_{it}}$ is the space-time interaction effect.
$\boldsymbol{\sigma_s, \sigma_u, \sigma_v, \sigma_\delta}$ are the standard deviation of the spatially-structured, spatially-unstructured, temporal and space-time random effect terms, respectively.
$\boldsymbol{1/\sigma_s^2}$, $\boldsymbol{1/\sigma_u^2}$, $\boldsymbol{1/\sigma_v^2}$ $\boldsymbol{1/\sigma_\delta^2}$ are the precision parameters

4.3.9. WinBUGS Implementation

The space-time separable and the three space-time inseparable models were fitted with Markov Chain Monte Carlo (MCMC) with different initial values for each model with a burn-in period of 116,000 iterations and thinning rate of 10 to obtain the posterior distributions of model parameters using WinBUGS software version 1.4. Mixings were observed using trace plots and autocorrelation plots. Convergence was evaluated by checking the Gelman–Rubin statistic[35] (Appendix 3 and Figure S4.2). After the burn-in period, a final run of 10,000 iterations for each chain was run to derive a final sample size of 20,000. The MCMC error of the model parameter estimates was <5% of the corresponding posterior standard deviations suggesting that the total 20,000 iterations, 10,000 from each of the two MCMC chains, are sufficient to provide a good approximation of the posterior distribution.

4.3.10. Model Selection

We assessed the Deviance information criteria (*DIC*) and the probability of direction (*pD)* values from the outputs to evaluate the goodness of fit for the four Bayesian hierarchical models and to select the best-fitted model (Meyer, 2016).

4.3.11. Spatial, Temporal and Spatiotemporal Relative Risks (RR)

The spatial, temporal and spatiotemporal relative risk estimates were obtained from the Model 3 (selected as the best-fitted model) outputs. Spatial relative risk ($RR_{Spatial} = \exp(S_i + U_i)$) (Haining & Li, 2020) for neighbourhood *i* represents the average COVID-19 incidence rate over 89 weeks in neighbourhood *i* compared to the average COVID-19 incidence rate in Toronto. A map of the posterior means of the spatial relative risk was created using the ArcGIS Desktop software 10.8.1. The temporal relative risk ($RR_{Temporal} = \exp(v_t)$) (Bie et al., 2021) at week *t* represents the average COVID-19 incidence rate for all neighbourhoods in week *i* compared to the average COVID-19 incidence rate of the entire study period. The posterior mean of the temporal relative risk with its corresponding 95% credible intervals was plotted using RStudio Software version 2021.09.0. The spatiotemporal effect term δ represents a change that cannot be reflected by spatial and temporal effects only. (Bie et al., 2021) The spatiotemporal relative risk ($RR_{SpatioTemporal} = \exp(\delta_{it})$),(Bie et al., 2021) represent the risk of COVID incidence rate in neighbourhood *i* and in time *t* compared to the overall incidence rate in entire study area and entire time period.

4.3.12. Joinpoint Regression

We used the joinpoint regression using the Joinpoint software version 4.9.0.0 (*Joinpoint Regression Program*, 2011; H. J. Kim et al., 2000), which uses the least-squares regression method to find the best-fit line from the temporal (weekly) pattern of the relative risk for COVID-19 derived from the Bayesian model. The joinpoint regression uses an algorithm that tests whether a multi-segmented line is a significantly better fit than a straight or less-segmented line. It involves fitting a series of joined straight lines on a log scale to the trends in the weekly

relative risk of COVID-19. Line segments are joined at points called joinpoints. Each joinpoint denotes a statistically significant (P = .05) change in trend. The significance test uses a Monte Carlo Permutation method to find the best-fit line for each segment.

4.3.13. Spatiotemporal Risk Trend Analysis

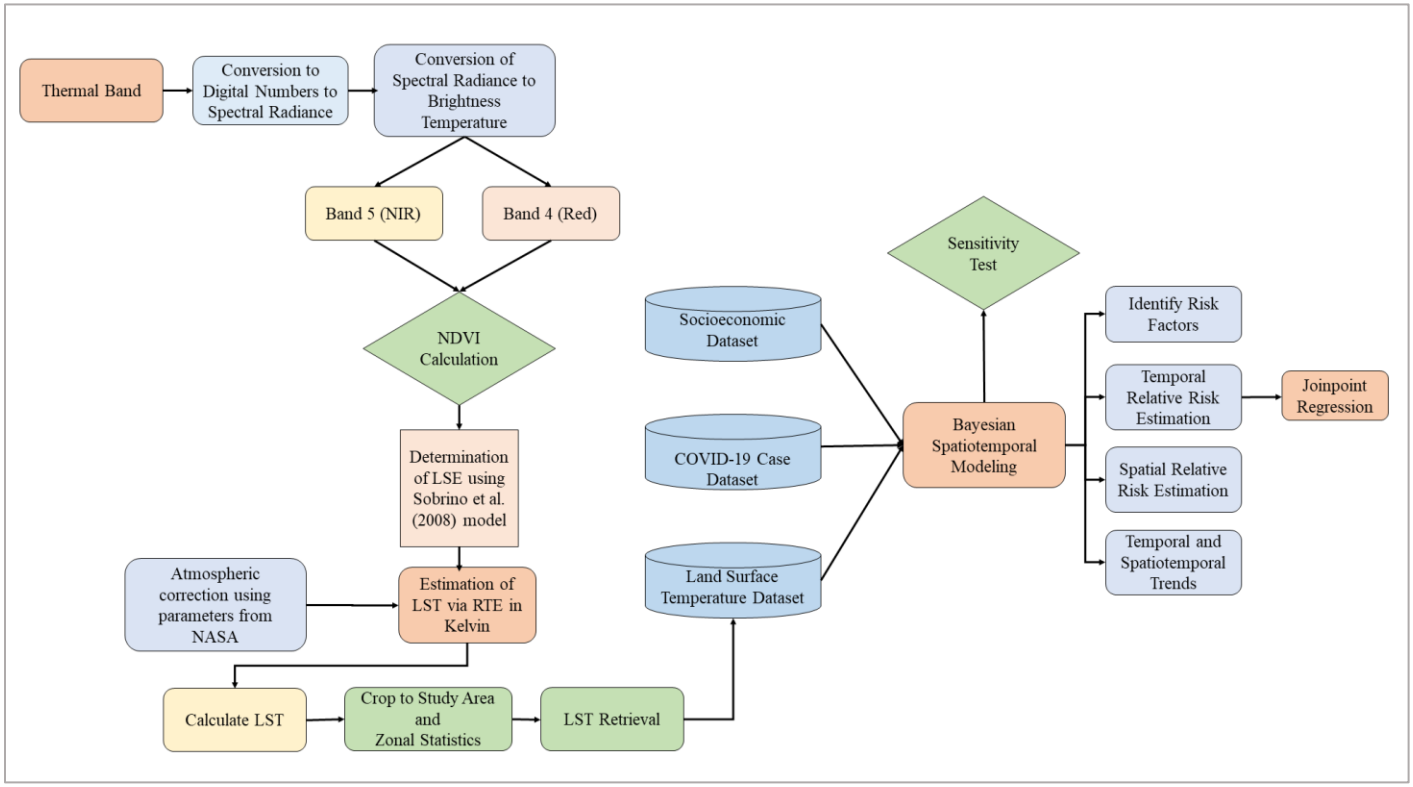
The spatiotemporal trend of the relative risk over time (increasing, decreasing, or stable) in the neighbourhoods was calculated in RStudio Software and mapped in ArcGIS. The neighbourhoods with a negative estimated coefficient were considered to have a decreasing trend over time, and the neighbourhoods with positive estimated coefficients were considered to have an increasing trend over time. The neighbourhoods with zero estimated coefficients were considered to have a stable trend over the study period.

4.3.14. Sensitivity Analysis

Given that there is no such thing as the true prior, two additional models with two alternate prior assumptions than the original final model (Model 3) were run to perform a sensitivity analysis to examine the final model results. The first model was run with hyperprior distributions of Gamma (0.005, 0.005) (Cai et al., 2013) on the precision parameters ($1/\boldsymbol{\sigma}_s^2$, $1/\boldsymbol{\sigma}_u^2$,$1/\boldsymbol{\sigma}_v^2$ $1/\boldsymbol{\sigma}_\delta^2$), and the second model was run with a uniform prior with a wider range (0.0001,1000) (Haining & Li, 2020) for the $\boldsymbol{\sigma}_s, \boldsymbol{\sigma}_u, \boldsymbol{\sigma}_v, \boldsymbol{\sigma}_\delta$ parameters (standard deviation of the spatially-structured, spatially-unstructured, temporal and space-time random effect term, respectively). The outputs from these two models with two different priors were compared to the outputs from the original model to ensure that our findings were not sensitive to the original hyperprior specification.

The methodological framework for our study is shown in a flow diagram (Figure 4.2).

Figure 4.2. The flow diagram of the methodological framework.



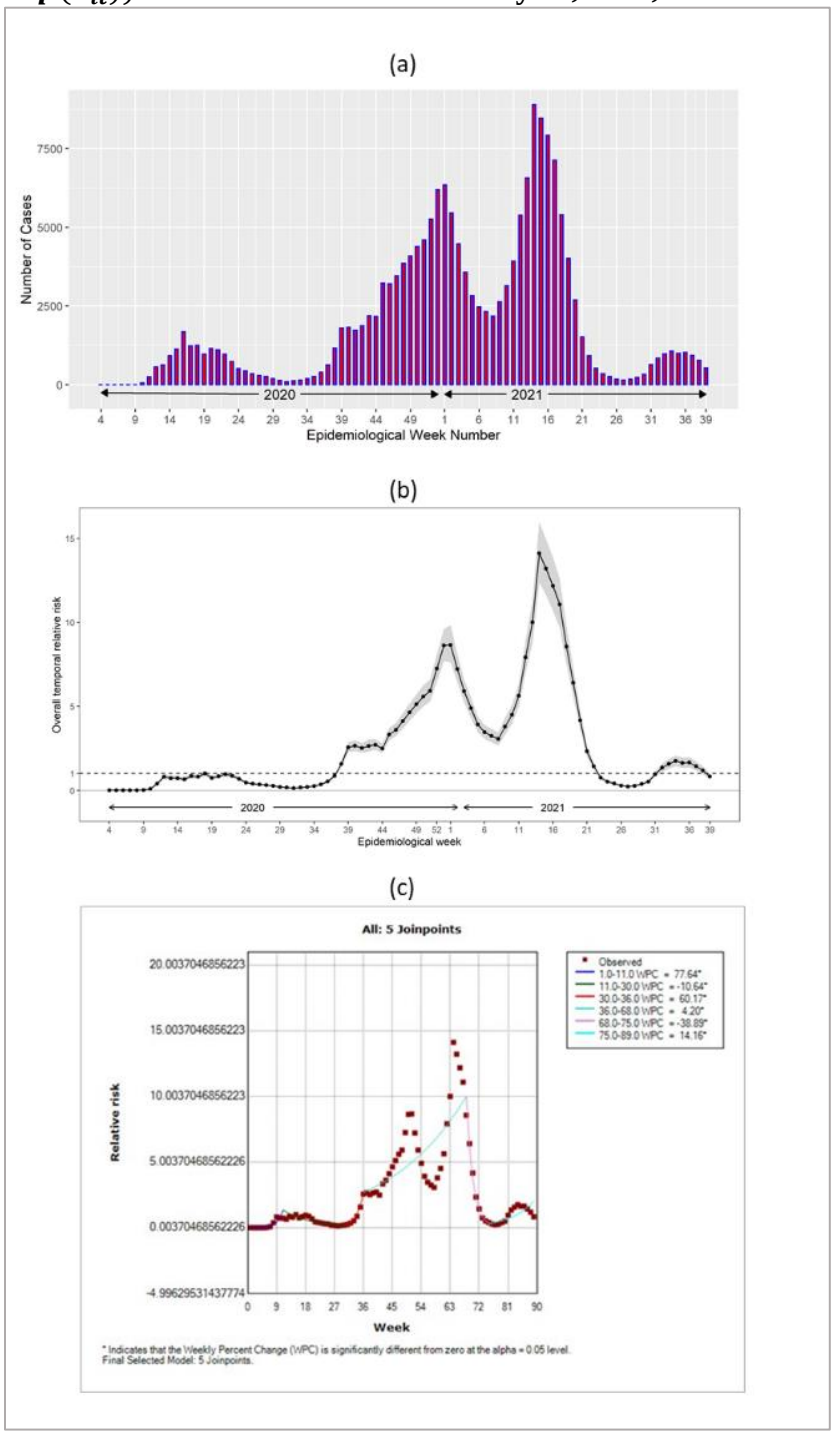
4.4 Result

4.4.1 Descriptive Statistics

We had 2,731,571 population in the study area based on the 2016 census data. A total of 179,072 (22,326 outbreak-associated and 156,746 sporadic) cases were reported in the 140 neighbourhoods during the study period. To avoid potential bias towards finding the high-risk clusters in the outbreak areas, our study excluded the outbreak-associated cases (12.4% of all cases), generally from healthcare (e.g., long-term care homes, hospitals) or congregated settings. We also excluded 2,423 (1.64%) cases due to missing neighbourhood information, leaving 154,323 sporadic cases for the analysis.

We observed a sharp increase in COVID-19 cases in late December of 2000, which declined in early January 2021 (Figure 4.3a). Again, a sharp increase in the cases was observed in late March (week 13) of 2021, which declined in late May of the same year. The highest number of cases (>7500 cases) were observed in April 2021.

Figure 4.3. (a) The weekly number of COVID-19 cases (excluding the outbreak cases) (b) Temporal Relative Risk (($RR_{Temporal} = exp(v_t)$) of COVID-19 in Toronto neighbourhoods (c) The spatiotemporal trend of the relative risks ($RR_{SpatioTemporal} = exp(\delta_{it})$) in Toronto between January 21, 2020, and October 2, 2021.



4.4.2. Model Selection

We fitted the data in four Bayesian Space-Time Hierarchical Models and compared the models (Table 4.3). The *DIC* values from the space-time inseparable models (Model 2-4) are much lower than the space-time separable model (Model 1), suggesting that the inseparable models (Model 2, 3 and 4) are better supported by the observed data, confirming the need to incorporate the space-time interaction component (Haining & Li, 2020). Finally, by comparing Model 2, 3 and Model 4, we found that Model 3 has the smallest *pD*, and *DIC* values, suggesting that Model 3 (space-time inseparable model with type II interaction) is the most parsimonious and therefore was selected as our final model.

Table 4.3 Comparison of the Four Bayesian Space-Time Models.

Bayesian Model	Space-Time Inseparable?	Space-Time Interaction Type	*pD*	*DIC*
Model 1	No	NA	246.697	66280.700
Model 2	Yes	Type I	5017.320	56711.300
Model 3	Yes	Type II	3231.73	55479.100
Model 4	Yes	Type III	4626.760	57043.700

4.4.3. Regression Outputs

Between January 21, 2020, and October 2, 2021, using the parameter estimates (mean $[\delta_{it}]$*1000) from the final model, the average COVID-19 rate per 1000 population per neighbourhood in Toronto is estimated to be 6.1 (95% CI: 6.0– 6.3). Table 4.4 reports the estimated relative risks by exponentially transforming the regression covariates associated with the three covariates from the Model 3 regression outputs. Based on the outputs, we observed that higher education was negatively associated (95% CI: 0.67-0.78), and LST (95% CI:1.01-1.17) was positively associated with COVID-19 incidence. An increase of one standard deviation in the higher education rate in a neighbourhood was associated with a 28% (95% CI: 22%-33%)

standard deviation decrease in the COVID-19 incidence rate. An increase of one standard deviation in average LST in a neighbourhood was associated with a 9% (95% CI:1%-17%) standard deviations increase in the COVID-19 incidence in Toronto. COVID-19 incidence was not found to be associated with the immigrants (95% CI: 0.98-1.05), as the 95% CI contains 1. Therefore, the percentage of immigrants does not appear to be a strong risk factor for explaining variability in the COVID-19 incidence in our study (Table 4.4).

Table 4.4. Estimated Relative Risks [$exp(\beta_k)$] and 95% CI.

Relate Risk (RR)	Posterior Estimates of Risk (95% CI)
$\mathbf{RR_{EDU}}$ (Higher Education)	0.72 (0.67-0.78)
$\mathbf{RR_{IMMI}}$ (Immigrants)	1.02 (0.98-1.05)
$\mathbf{RR_{LST}}$ (LST)	1.09 (1.01-1.17)

Temporal Relative Risk

Figure 4.3b presents the posterior mean and 95% uncertainty band of the temporal relative risk (($RR_{Temporal} = \exp(v_t)$) (Bie et al., 2021) during the study period, which shows that the highest risk (RR>9) was observed between March 14, 2021, and April 17, 2021. A total of 46 (51%) weeks out of the total 89 weeks experienced a relative risk of less than one during our study period.

Joinpoint Regression Results

Figure 4.3c represents the temporal patterns of relative risk for COVID-19 from January 21, 2020, to October 2, 2021 (89 consecutive weeks). The line displays five joinpoints (6 line segments or trends), indicating a significant change in the relative risk six times during the study period. For instance, the relative risk of COVID-19 was increased by 77% per week from the

beginning to the 11th week. The risk was then decreased by 11% by the 30th week, then was increased by 60% until the 36th week, then increased only by 4% until the 69th week, then it decreased by 39% until the 75th week, and it again increased by 14% by the end of the study period.

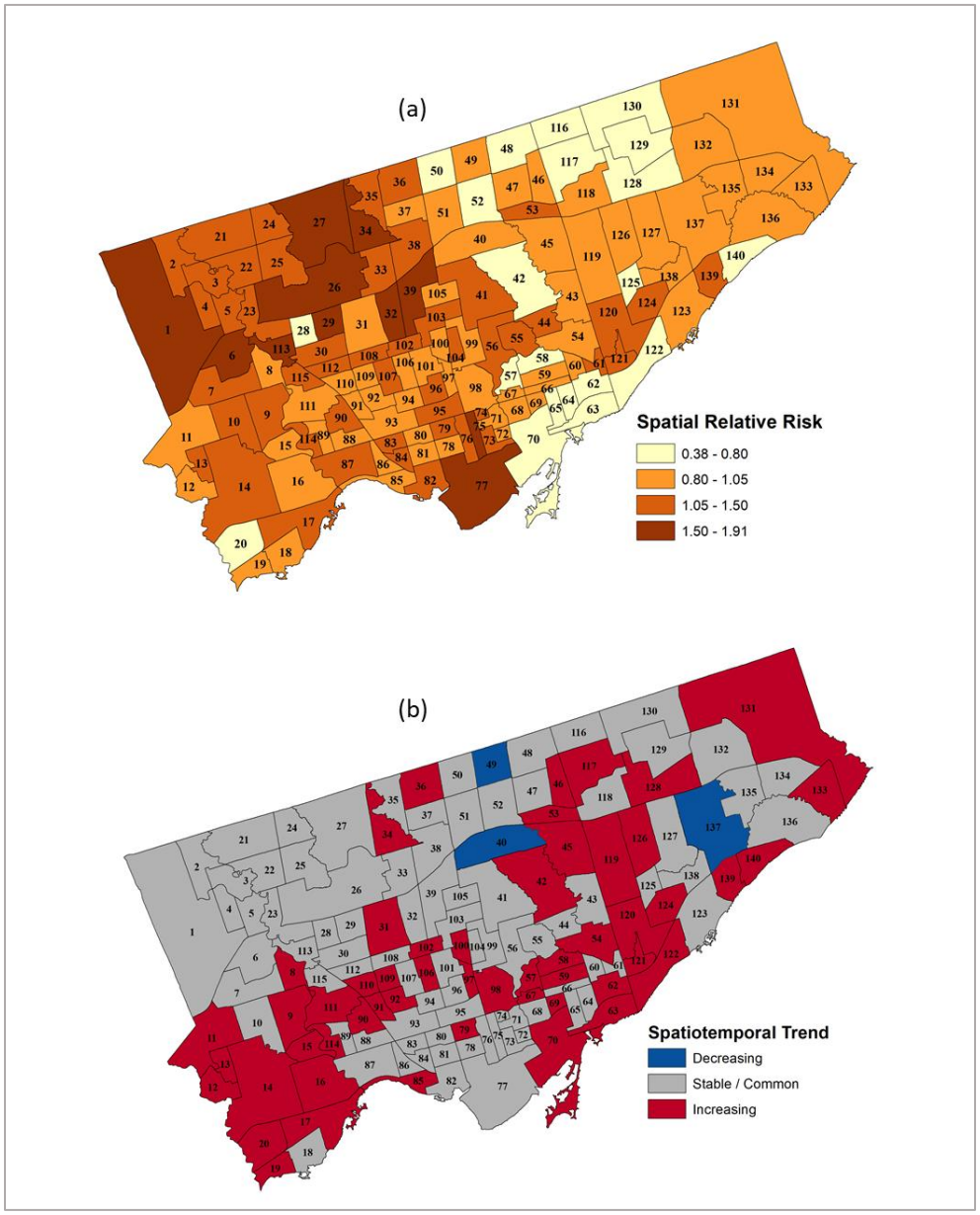
Spatial Relative Risk

The posterior means of the relative risks (spatial) of COVID-19 in the Toronto neighbourhoods are presented in Figure 4.4a. A high risk (RR>1.5) was observed in northwestern and southern (Neighbourhood # 77) Toronto. A moderate level of risk (RR>1.05) was observed in different neighbourhoods in western Toronto. Eastern Toronto mostly experienced a low risk of COVID-19.

Spatiotemporal Relative Risk Trends

While evaluating the trends of the spatiotemporal relative risks ($RR_{SpatioTemporal} = \exp(\delta_{it})$) (Bie et al., 2021), we observed that only three neighbourhoods (Neighbourhoods # 40,137 and 49) had a decreasing trend of the relative risk during the study period (Figure 4.4b). A total of 84 (60%) neighbourhoods had a stable trend, and 53 (37%) neighbourhoods had an increasing trend in the study area.

Figure 4.4. (a) A map of the estimated overall spatial pattern based on the posterior means of spatial relative risks ($RR_{Spatial} = exp(S_i + U_i)$) (b) The spatiotemporal trend of the relative risks ($RR_{SpatioTemporal} = exp(\delta_{it})$) in Toronto between January 21, 2020, and October 2, 2021.



Note: The numbers inside the neighbourhood represent the neighbourhood identification number.

Sensitivity Test Results

The sensitivity test models with Gamma (0.005, 0.005) (Cai et al., 2013) priors for the precision parameters and uniform (0.0001,1000) (Haining & Li, 2020) priors for standard deviation parameters are equally appropriate, giving almost identical posterior distributions of the parameters and model fit (DIC values) compared to outputs of the original model (Model 3), presented in Appendix 4 in the supplementary materials. Therefore, we can conclude that an apparently innocuous uniform prior that we have used in our final model is not introducing substantial information into the model fitting.

We mapped the spatial patterns of COVID-19 cases during the post-study period (October 3-31, 2021) to visually compare the spatial patterns of risk identified from the Bayesian spatiotemporal model (Figure S4.3 in Appendix 5 in Supplementary materials). As observed, the spatial risk obtained from the Bayesian model was higher in the northwestern and southern neighbourhoods of Toronto. In these neighbourhoods, a higher number of cases were also reported during the post-study period. The map of the spatiotemporal trend showed an increasing pattern in the eastern, southwestern and central neighbourhoods. These neighbourhoods experienced a higher number of cases during the post-study period. In Northwestern Toronto, the trend was stable but observed a higher spatial risk, indicating that the area remained at higher risk throughout the study period. This region also showed a higher number of cases during the post-study period.

4.5 Discussion

In our study, we observed spatial, temporal, and spatiotemporal trends of COVID-19 in Toronto and identified the key factors associated with the transmission of the disease. Overall, the trend and transmission patterns of the disease were heterogeneous over space and time. Only three neighbourhoods experienced a decreasing spatiotemporal risk trend in the area. Most neighbourhoods experienced either stable or increasing spatiotemporal risk during the study period. We also observed several high-risk neighbourhoods in the western and southern parts of Toronto, and the risk in those neighbourhoods remained constant throughout the study period. Since educational status and LST were associated with the risk, we believe these factors might have influenced to remain those areas at high risk throughout the study period.

The temporal risk was particularly high in the early spring of 2021, suggesting that the temperature during the season could have influenced the transmission of the disease in that part of the country. However, some other factors, such as an increase in mobility or travel patterns during the long weekend in March 2021 (Reuters et al., 2021), could also influence the increase of the disease in spring. The disease risk remained high for more than half of the study period. We observed a lower risk during the early summer (June-July) of 2020 and 2021, which could be due to an increased time spent in outdoor settings leading to a decreased COVID-19 risk (Brown et al., 2021). We identified significant changes in the risk of COVID-19 six times across the study period with varying trends. These changes suggest that the temporal trend of the epidemic of COVID-19 is different from other coronavirus diseases, such as the SARS (severe acute respiratory syndrome) epidemic in 2003, where the number of reported SARS cases has

increased exponentially over time, and the outbreak lasted approximately six months (Boulos, 2004; CDC, 2021; Galvani et al., 2003; X. Zhang et al., 2020).

The findings of the positive association of COVID-19 with LST in our study are consistent with the findings of an earlier study. One reason behind this could be that high LST favours the coronavirus. Furthermore, Toronto is the most densely populated urban city in Canada. A recent study conducted in Phoenix, Arizona and Dallas, Texas, in the United States by Moss and Kar (Moss & Kar, 2020) concluded that the urban areas that are susceptible to a high Urban Heat Index, as measured by LST, are primarily occupied by vulnerable population groups. Therefore, it is possible that an excess vulnerable population in an area creates a higher LST in the neighbourhood, resulting in an increase in the risk of COVID-19 in our study area.

Our study also finds that the higher education rate has a negative association with COVID-19, which is in line with the findings from earlier studies (Abedi et al., 2020; Cordes & Castro, 2020; Goutte et al., 2020; Y. Wu et al., 2020). It is likely that people with higher education may understand and follow public health messages as well as have the option to work remotely and maintain social distancing, resulting in lower incidences of COVID-19 in areas with a higher number of educated people. Additionally, a study by Mondal et al. in 2021 found that higher education levels were associated with a higher likelihood of vaccine acceptance (Mondal et al., 2021), suggesting that intervention with health education may play a key role in fighting this pandemic. Various public health programs, such as COVID-19 awareness and health education programs in neighbourhoods with low education, may also help reduce the fast transmission rates in those neighbourhoods.

In our study, we used higher spatial and temporal resolution satellite images to extract LST. We also used atmospheric corrections methods on these images by adopting Sobrino et al. in 2008's Land Surface Emissivity (LSE) model (Sobrino et al., 2008), which provided a high estimation accuracy. Our findings may advocate for maintaining disease surveillance and planning for an effective public health program. Most of the earlier studies (Briz-Redón & Serrano-Aroca, 2020; Y. Ma et al., 2020; H. Qi et al., 2020; Rashed et al., 2020; Runkle et al., 2020; Tobías & Molina, 2020) explored the relationship between ambient temperature or LST and COVID-19 at a broader spatial scale (provincial or state) and used a shorter period with a very limited number of images to extract LST without any atmospheric corrections (Das et al., 2020). We believe that our study filled the gap in the existing literature by using higher spatial and temporal resolution satellite imageries at a local spatial scale, which is more spatially representative and may have provided a more accurate estimate due to the use of the atmospherically corrected data on LST.

The validation results of the spatiotemporal patterns of risk using the data of the post-study period suggest that the Bayesian model could predict spatial patterns of risk for COVID-19 in our study area. Therefore, the findings of our study can be useful for increasing awareness of the disease and preparing public health interventions aimed at targeted prevention and control of COVID-19. Given limited resources available, efforts could focus on the high-risk neighbourhoods, as observed in our study.

Our study has several limitations. First, we could not find data on air pollution or human mobility patterns at the neighbourhood level, which could be important contributors to influencing the COVID-19 incidence. Second, we had 1.6% of cases with missing

neighbourhood information and 12.4% cases were outbreak cases, which were not included in our analysis. Additionally, COVID-19 is often asymptomatic, under-reported (P. Wang et al., 2022) or lacks accurate information on the onset of COVID-19, limiting the capacity of the analysis. However, the Bayesian spatiotemporal hierarchical models allowed us to compensate for the missing/unobserved covariates or missing data by incorporating the structured, unstructured random effects into the model (Knorr-Held, 2000). In particular, the Type II space-time interaction in our final model implied that the missing covariates have smoothly varying structures through time and have no structure over space since they are highly localized in their effect on the outcome (Knorr-Held, 2000). Third, our study has an ecological study design where the data were aggregated at the neighbourhood level, which may create issues such as ecological fallacy (Sedgwick, 2015). Therefore, these results cannot be interpreted at the individual level.

Despite these limitations, our study, due to its strong versatility and complex hierarchical modelling, is still convincing and has provided important information that may improve our understanding of the transmission patterns of COVID-19 and the associated risk factors. Also, our model is superior to the frequentist method that is more frequently used, as the Bayesian approach allowed us to compensate for the missing covariates in the models in identifying spatial patterns of risk. Therefore, we believe that using the Bayesian spatiotemporal model and the long-time series satellite-derived environmental data for modelling disease transmission have advanced our understanding of the disease risk in space.

4.6. Conclusion

Several conclusions can emerge from our study. First, the Bayesian analysis has shown that Bayesian regression with spatial (structured and unstructured), temporal and spatiotemporal

random effects provided an effective framework for understanding COVID-19 disease transmission. Second, the spatiotemporal risk remained high for the entire study period and constantly high for the high-risk neighbourhoods. However, the temporal risk fluctuated over time in the study area. Third, higher education and LST played an important role in predicting COVID-19 incidence. Therefore, it is important to take those factors into account while planning intervention strategies. Fourth, the framework presented in this study may help make an early warning system for COVID-19 incidence and assist public health authorities in controlling and preventing outbreaks of similar diseases. Finally, the methodological framework applied here can also be used in other small area-level studies on infectious diseases.

Chapter 5 Manuscript 3: Modelling the spatiotemporal spread of COVID-19 outbreaks and prioritization of the risk areas in Toronto, Canada

5.1 Summary

Modelling the spatiotemporal spread of a highly transmissible disease is challenging. We developed a novel spatiotemporal spread model, and the neighbourhood-level data of COVID-19 in Toronto was fitted into the model to visualize the spread of the disease in the study area within two weeks of the onset of first outbreaks from index neighbourhood to its first-order neighbourhoods (called dispersed neighbourhoods). We also model the data to classify hotspots based on the overall incidence rate and persistence of the cases during the study period. The spatiotemporal spread model shows that the disease spread to 1-4 neighbourhoods bordering the index neighbourhood within two weeks. Some dispersed neighbourhoods became index neighbourhoods and further spread the disease to their nearby neighbourhoods. Most of the sources of infection in the dispersed neighbourhood were households and communities (49%), and after excluding the healthcare institutions (40%), it becomes 82%, suggesting the expansion of transmission was from close contacts. The classification of hotspots informs high-priority areas concentrated in the northwestern and northeastern parts of Toronto. The spatiotemporal spread model along with the hotspot classification approach, could be useful for a deeper understanding of spatiotemporal dynamics of infectious diseases and planning for an effective mitigation strategy where local-level spatially enabled data are available.

5.2. Introduction

The novel Coronavirus-2019 (COVID-19), first reported in Wuhan, Hubei province, China, in 2019, is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that causes respiratory illness in humans (Marchand-Senécal et al., 2020). As of July 15, 2022, World Health Organization (WHO) reported over 556.9 million confirmed COVID-19 cases and 6.4 million deaths globally (WHO, 2021). The disease is highly contagious, often without symptoms that seriously challenge the health care system (Smereka & Szarpak, 2020). COVID-19 has also greatly impacted Canada, and the federal and provincial governments have adopted several control strategies to control the spread (Canada, 2022). Due to its fast infection transmission rates, large urban cities such as Toronto are often more affected than rural areas as they are more densely populated, making it hard to control and maintain social distance in the urban areas(Paul et al., 2020). The recommendations of policies and research programs to combat an epidemic are generally based on the perceived standardized definitions of an outbreak. Infections can spread among humans by contacts within the household, community or by random contacts among the general population (Balcan et al., 2010). Timely detection of the spatial spread of infectious disease outbreaks is important for control efforts and can provide clues to the important risk factors (Chowell & Rothenberg, 2018a; Steele et al., 2020).

Infectious diseases rarely exhibit simple dynamics (Harapan et al., 2020). COVID-19 is an infectious disease with spatial dimensions that have helped researchers and scientists understand the disease epidemiology of this phenomenon. One of the most common ways to observe outbreak dynamics is the use of visualization techniques that allow us to easily identify spatial patterns and select areas of interest for further studies. Identifying the spatial patterns of the outbreaks and further analysis in those areas can help researchers understand the disease

dynamics, which may help in policy-making for controlling the outbreaks (Carroll et al., 2014). These types of maps are often effective in locating an outbreak's origin and depicting outbreak progression over time.

During an epidemic, observing the movement of outbreaks across neighbourhoods in space and time may allow identifying possible sources and directions of the outbreaks. Earlier studies that mapped the spatiotemporal spread of infectious disease outbreaks mostly involved comparing multiple maps to compare the outbreak shifting visually. For instance, a few studies evaluated the spread of COVID-19 over space and time by visually comparing multiple maps to observe the spatial distribution of new or cumulative cases using multiple time points during the early stages of the pandemic (Cuadros et al., 2020; Y. Feng et al., 2020; Gao et al., 2021; H. Yu et al., 2021). Standard deviation ellipse (SDE), a popular spatial statistical method, is often implemented in outbreak studies to model the spread of disease over time to analyze the spatiotemporal spread of an infectious disease by using the mean centre of the spatial distribution of the geographical elements as the centre and calculates the standard deviation of SDE in the *X* and Y axis (Gesler, 1986; W. Li et al., 2022; X. Wang et al., 2022). Multiple days of SDE maps are often compared to observe the differences in directions between multiple time periods. This method analyzes directional distributions of an outbreak; however, do not provide insights into the magnitude of these transmissions from index to dispersed neighbourhoods.

A study by Gianquintieri et al. in Lombardy, Italy, used choropleth mapping to display the beginning of COVID-19 onset from emergency calls while not including details about the spread mechanism (Gianquintieri et al., 2020). Spread patterns were modelled in other infectious diseases as well, such as the 2009 pandemic influenza outbreaks. A transmission model by Gog

et al. showed the spatial spread of the 2009 pandemic influenza but lacked further details about the starting point and the direction of movement of the outbreaks (Gog et al., 2014). Reyes et al. observed the spatiotemporal spread patterns of the 1918 influenza pandemic in British India by using a travel network and a likelihood approach to predict the patterns of pandemic spatial spread. The output map shows the travel network connectivity between districts using nodes across British India, and the nodes with the high excess mortalities are marked in red to explain the spatial spread (Reyes et al., 2018). Another study by daCosta et al. visually examines the spatiotemporal diffusion of influenza A (H1N1), identifies the starting point of an epidemic, and further uses multiple maps of the time of onset and the correlation between effective distance trees across different spatial units (da Costa et al., 2018). However, these existing spatial transmission or spread models often overlook the temporal structures, the direction and the magnitude of the transmission processes within the surrounding neighbourhood networks and primarily compare disease hotspot maps from multiple time points to visualize the dispersion processes. Since the transmission process of the disease may vary geographically and in time, it is important to investigate the transmission dynamics of the diseases using a novel model at a local scale. These types of maps can be effective in locating the origin of an outbreak and depicting outbreak progression over time.

Displaying the outbreaks as they are happening and possible spread to the nearby neighbourhoods can provide deeper insights, allow a more effective response to outbreaks, and provide the ability to understand the intricate details and plan on how to respond to them. In this study, we developed a novel approach for visualizing the spatial spread of the COVID-19 outbreaks at a neighbourhood level that possibly suggests local transmission patterns and the

direction of movements of a highly infectious disease in an urban context. The spatial spread of the disease in the neighbourhoods was also classified into priority-based hotspots for facilitating the intervention program for controlling the spread of the disease under resource-constraint situations.

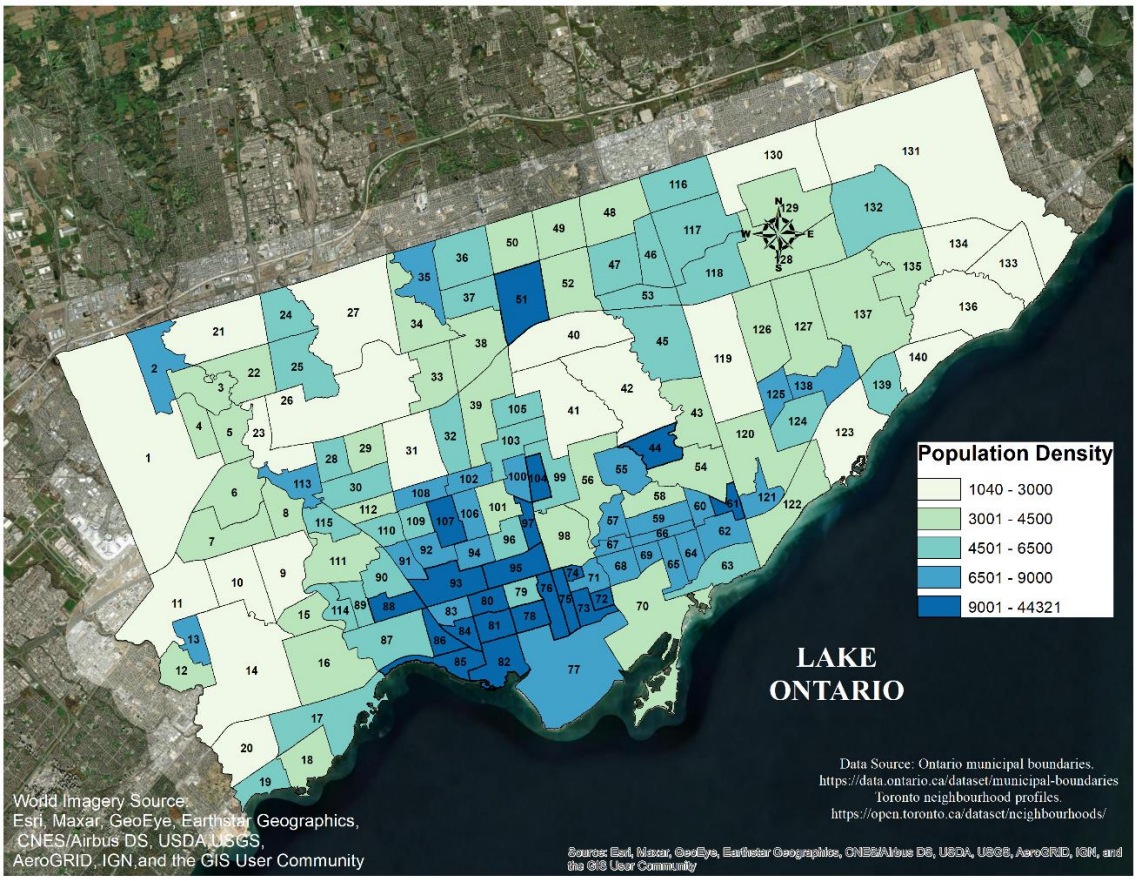
5.3. Methods

5.3.1. Study Design, Study Area and Population

This study was a retrospective observational study using existing data collected at the neighbourhood level in Toronto (Figure 5.1). The city of Toronto is on the northern and western shorelines of Lake Ontario in Canada, with a land area size of 630 square kilometres. Toronto is the largest city in Canada and the fourth largest city in North America, with a high population density (Figure 5.1) of 4,692 persons per square kilometre (City of Toronto, 2021a). Toronto Public Health has reported the highest total number of COVID-19 cases (342,301 cases as of July 15, 2022) and the second-highest incidence rate among all the health regions in the nation (City of Toronto, 2020a; Government of Canada, 2020). Over 4,322 deaths have been attributed to COVID-19 in Toronto since March 24, 2020 (City of Toronto, 2020a).

Our study area consists of 140 geographically distinct social planning neighbourhoods with detailed demographic information for each neighbourhood, as prepared by the city's Social Policy Analysis & Research unit from the data of the 2016 Statistics Canada Census (updated every five years) (City of Toronto, 2020b). The neighbourhoods were defined based on Statistics Canada census tracts for the purposes of statistical reporting (Statistics Canada, 2021). The average population size in a neighbourhood is 19,511 persons (minimum = 6,577, maximum = 65,913) as per the 2016 census.

Figure 5.1. The Study Area in Toronto, Ontario, Canada.



5.3.2. Data

The total number of COVID-19 cases per neighbourhood from January 2020 to June 2021 was retrieved from the city of Toronto's open data portal site (City of Toronto, 2020a). The dataset was collected by Toronto Public Health that contains demographic, geographic, and severity information for all confirmed and probable cases. We used the date of episodes as the time of case occurrences of the disease. In the surveillance database, the date of the episode was chosen as the first day that COVID-19 symptoms, the date of specimen collection, or the date of reporting, whichever is the earliest (City of Toronto, 2020a).

The dataset includes the source of infection for each case (City of Toronto, 2022a), which was determined by the Toronto public health authority by assessing the most likely source of infection out of various risk factors such as travel or contact. The common sources of infection included household or close contact such as family members or roommates with confirmed or probable COVID-19 cases, outbreaks in congregate settings such as shelters, correctional facilities, group homes or hostels or healthcare institutions such as long-term care homes, retirement homes, hospitals and chronic care hospitals. Other sources of infection include travel (travelling outside of Ontario in the 14-day prior to their symptom onset or test date), outbreaks in other settings, including workplaces, schools or daycares, and the community (individuals who did not travel outside of Ontario and did not identify being close contact with a COVID-19 case and were not part of a known confirmed COVID-19 outbreak) (City of Toronto, 2022a). Approximately 26% of the total COVID-19 cases had no information on the source of infection.

The population data per neighbourhood were collected from the city of Toronto neighbourhood profiles based on the 2016 census. According to the 2016 census used in our study, the total population of the study area is 2,731,571. The full dataset can be found at the city of Toronto's open data portal site (City of Toronto, 2020b). The neighbourhood-level administrative boundary shapefile for Toronto was extracted from the same site (City of Toronto, 2020b). We aggregated daily COVID-19 cases by the epidemiological week, resulting in a total of 76 epidemiological weeks for each of the 140 neighbourhoods in Toronto. Since this study observes the overall burden of the disease, we included both sporadic and outbreak cases in this analysis.

5.3.3. Research Methods

First, we created the spatiotemporal spread model that identified the index and the dispersed neighbourhoods of COVID-19 outbreaks. We have also analyzed the sources of infections in the dispersed neighbourhoods. Second, the priority-based hotspots were defined in the hotspot classification model based on the two epidemiological parameters: incidence rate and disease persistence in each neighbourhood. Sensitivity analyses were also performed for the spatiotemporal spread model and the hotspot classification model.

5.3.3.1. The Spatiotemporal Spread Model

We developed a spatiotemporal spread model to visualize the spatial dynamics of COVID-19 across neighbourhoods. At first, the model checked the number of cases per week for each neighbourhood. If the observed number of cases in a week in a neighbourhood was more than the weekly average number of cases of all time points and all neighbourhoods, it was considered an outbreak event in the neighbourhood, a commonly used approach in infectious disease surveillance . The model then checked the time of the first outbreak in each neighbourhood. Only the first outbreak event in a neighbourhood and the first-order neighbourhood networks sharing a common border of non-zero length (queen contiguity) were considered in evaluating the transmission dynamics in space and time. The index neighbourhoods were considered as the starting point of the dispersion event, and the dispersed neighbourhoods are the neighbourhoods where the first outbreaks happened within two weeks of the outbreak in the index neighbourhood.

We used the two-week temporal period, as the World Health Organization and national and international public health sectors have widely accepted the 14-day incubation period for COVID-19 infection spread (Quilty et al., 2021). Also, the two-week time was chosen based on the assumption that further spread would likely be decreased two weeks after the onset of the event. If we get a dispersed neighbourhood, we connect the neighbourhood with the index neighbourhood to show the potential spatial spread of the disease. If there was an outbreak event in a neighbourhood but not dispersed to any neighbourhood, then the neighbourhood was not considered an index neighbourhood. The index and dispersed neighbourhoods were, therefore, determined by

$$\boldsymbol{d = w_{ij}(o_j - o_i)} \qquad (1)$$

where o_i is the start week of the 1st outbreak event in neighbourhood i,

o_j is the start week of the 1st outbreak event in neighbourhood j,

w_{ij} is a measure of adjacency between neighbourhood i and j and is defined as (1 if i and j are adjacent; 0 otherwise).

When d is ≤ 2 (i.e., the difference between the start week of an outbreak between the two adjacent neighbourhoods is within two weeks), then i is the index neighbourhood, and j is the dispersed neighbourhood of the outbreak.

Since disease outbreaks definitions are implicitly variable (Brady et al., 2015), we performed a sensitivity test using observed cases in a week that is greater than the weekly median number of cases but keeping the temporal period to two weeks to observe how the outcomes of our spatiotemporal spread model fluctuate when using the other definition of an

outbreak. We also performed another sensitivity test by changing the temporal period from two weeks to one week but keeping the outbreak definition of the main analysis (observed cases in a week is greater than the weekly mean number of cases) to observe the differences in the results of the spatiotemporal model when using a shorter incubation period.

We have further divided the total study period into three temporal periods based on the first 3 COVID-19 waves: Wave 1 (January 23 – July 17, 2020), Wave 2 (July 18, 2020 – March 4, 2021) and Wave 3 (March 5 – July 31, 2022) and ran the spatiotemporal spread model for each of this period to understand the variations of the spatial spread of outbreaks across the three waves.

5.3.3.2. Classification of the Hotspots

Priority-based hotspots have previously been used in various infectious disease mapping (Debes et al., 2021; Ngwa et al., 2021). In this study, we used two epidemiological parameters to define the priority-based hotspots: i) overall incidence rate and ii) disease persistence, i.e., number of weeks with cases. The incidence rate for each neighbourhood was calculated using the total number of cases during the entire study period divided by the 2016 population (the latest available dataset) multiplied by 1000. The disease persistence was the total number of weeks with at least a case in the week (not necessarily the consecutive week). Based on these two epidemiological parameters used elsewhere (Dom et al., 2010; Mwaba et al., 2020), the neighbourhoods were classified into three categories of priority levels, as shown in Table 5.1 The cut-offs for the thresholds were determined based on the distribution of the data. The high priority hotspots consist of a higher incidence rate and longer persistence; the medium priority levels hotspots experience a lower incidence rate, but longer persistence or higher incidence rate

but shorter persistence, and the low priority hotspots consist of a lower incidence rate and shorter persistence.

Table 5.1. Thresholds of the epidemiological parameters applied for the priority-based mappings.

Priority type	Interpretation	Incidence Rate (70^{th} percentile value)	Disease Persistence (70^{th} percentile value)
P1	High Priority	> 76.29 cases/1000 persons	>65 weeks
P2	Medium Priority	≤ 76.29 cases/1000 persons	>65 weeks
P3	Medium Priority	> 76.29 cases/1000 persons	≤65 weeks
P4	Low Priority	≤ 76.29 cases/1000 persons	≤65 weeks

Since prioritizing the hotspots for intervention depends on the availability of the resources, one may choose the cut-off of the epidemiological parameters based on the need and logistical support. Keep that in mind, we performed sensitivity analyses using different cut-offs of the epidemiological parameters to demonstrate how one can go with the decision-making for prioritizing the hotspots. These cut-offs were: i) 50^{th} percentile values for both incidence rate and persistence (50 cases/1000 persons, 50 weeks), ii) 70^{th} percentile value (76.29 cases/1000 persons) for incidence rate and 50^{th} percentile value (61 weeks) for the persistence, and iii) 50^{th} percentile value for incidence rate (50.37 cases/1000 persons) and 70th percentile (65 weeks) value for the persistence.

5.3.3.3. Software

The analyses and mapping of this study were conducted using R Studio Version 1.14.1103 and ArcGIS Desktop Version 10.8.1, respectively. We used the following R packages: lubridate, sqldf, data.table, sqldf, rgdal, spdep for writing the codes for the spatial spread model, i.e., identifying index and dispersed neighbourhoods, and the resulting maps were produced using

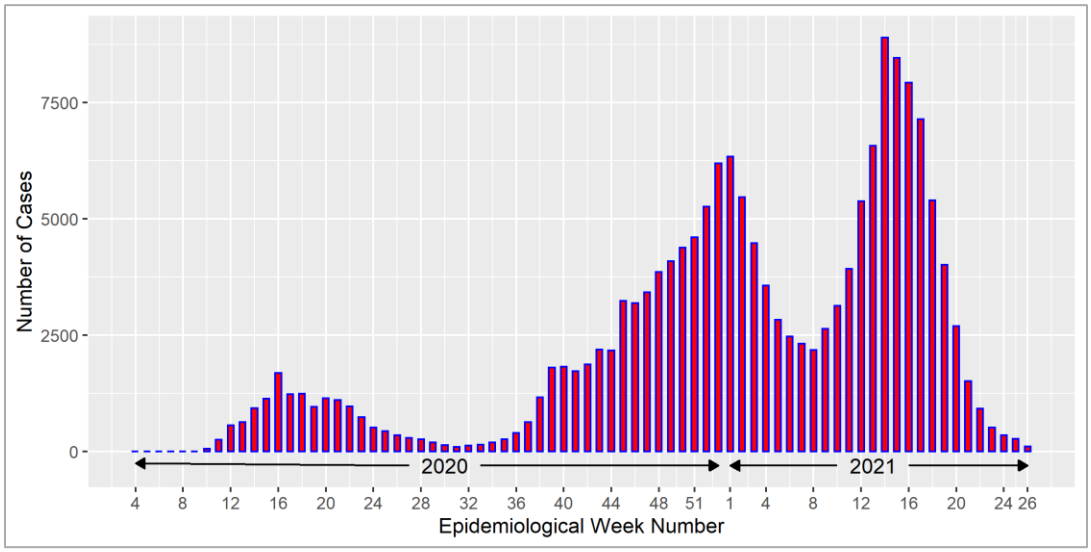
ArcGIS Desktop. The paths of the outbreaks were created in ArcGIS Desktop by connecting the centres of the index and dispersed neighbourhoods. For the priority-based hotspot maps, we used the following R packages: lubridate, sqldf, dplyr, magrittr and grid. The model outcomes were imported into ArcGIS Desktop to create the maps. We used ggplot2 and ggrepel packages in R to create the figures.

5.4. Results

5.4.1. Descriptive Statistics

During the surveillance period, a total of 169,985 COVID-19 cases were reported. We omitted 2,736 (1.6%) cases due to missing neighbourhood names, resulting in 167,249 (146,177 sporadic and 21,072 outbreak-associated) cases for analysis. An average of 19 cases were observed per week in a neighbourhood. The weekly number of cases shows that COVID-19 was reported throughout the study period (Figure 5.2). There was a sharp increase of cases in week 16 of 2020, which declined in the following weeks. In week 45 of November 2020, the number of cases increased markedly until the first week of 2021. Another sharp increase was observed in week 9 (March) and continued until April 14-21, 2021. The highest number of cases (>8000 cases) were observed in the week of March 14-21, 2021 (Figure 5.2).

Figure 5.2. The weekly total number of COVID-19 cases in Toronto between January 2020 and June 2021.

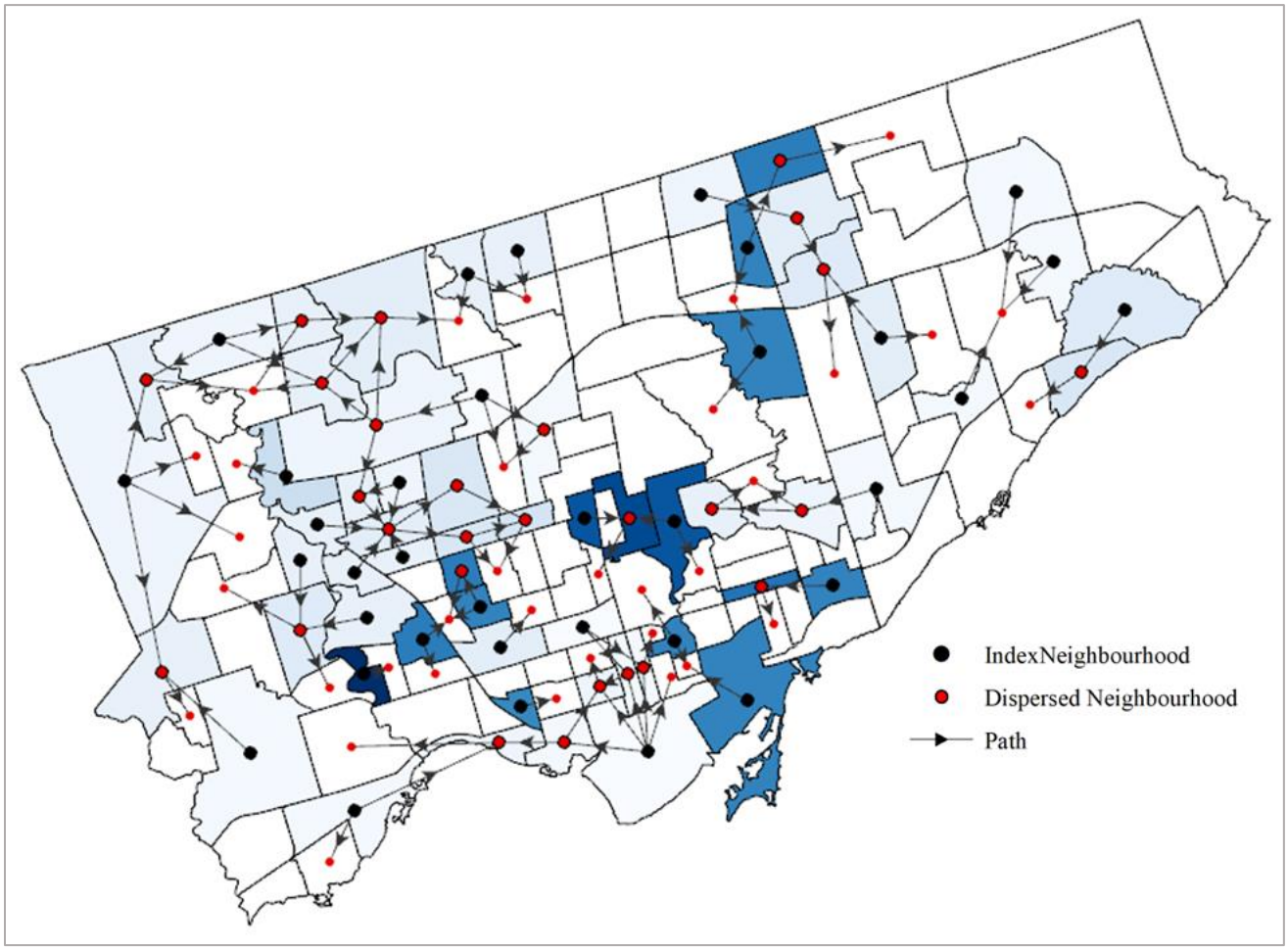


5.4.2. The Spatiotemporal Spread Model

The resulting map from our spatiotemporal model is shown in Fig 5.3, where the black dots represent the index neighbourhoods (62 (44%) neighbourhoods), which dispersed the outbreaks to their bordering neighbourhood(s). The red dots represent dispersed neighbourhoods (61 (43%) neighbourhoods), implying that the outbreak was possibly dispersed from the index neighbourhoods, as indicated by the path. In some cases, the dispersed neighbourhoods became the index neighbourhoods (27 (19%) neighbourhoods), potentially spreading the outbreak to nearby neighbourhoods. The earlier onset of outbreaks was observed in the western and southern parts of Toronto (shown in lighter blue gradient colours), whereas the later onsets were seen in the central parts of Toronto (shown in darker blue gradient colours). In the northwestern, northeastern, and southern parts of Toronto, several dispersed neighbourhoods became index neighbourhoods and further spread the disease to its nearby neighbourhoods. A maximum of 4 dispersed neighbourhoods was observed bordering the index neighbourhood. During the study

period, the last dispersion of the COVID-19 outbreak occurred between 18 and 24 April 2021. No dispersion events were observed in 90 (64%) neighbourhoods within two weeks of the onset of the outbreak in these neighbourhoods. Four neighbourhoods in the center of Toronto never experienced an outbreak event, as per our definition of the outbreak.

Figure 5.3. Spatiotemporal spread dynamics of the COVID-19 outbreaks in Toronto between January 2020 and July 2021.



Note: The lighter the colour the earlier the onset of the first outbreak in the neighbourhood.

Some dispersed neighbourhoods, particularly in northwestern, southern and central Toronto, had multiple index neighbourhoods, as shown in Fig 5.3. Since these neighbourhoods were shown to have dispersed infections from multiple index neighbourhoods, we also created a

distance-based analysis where only one neighbourhood was determined as the index neighbourhood for a dispersed neighbourhood (Fig S3). If a dispersed neighbourhood had multiple index neighbourhoods, only the neighbourhood that was the closest (based on the linear distance from the centers of the neighbourhoods) was considered the index neighbourhood.

5.4.2.1. Sensitivity Analyses of the Spatiotemporal Spread Model

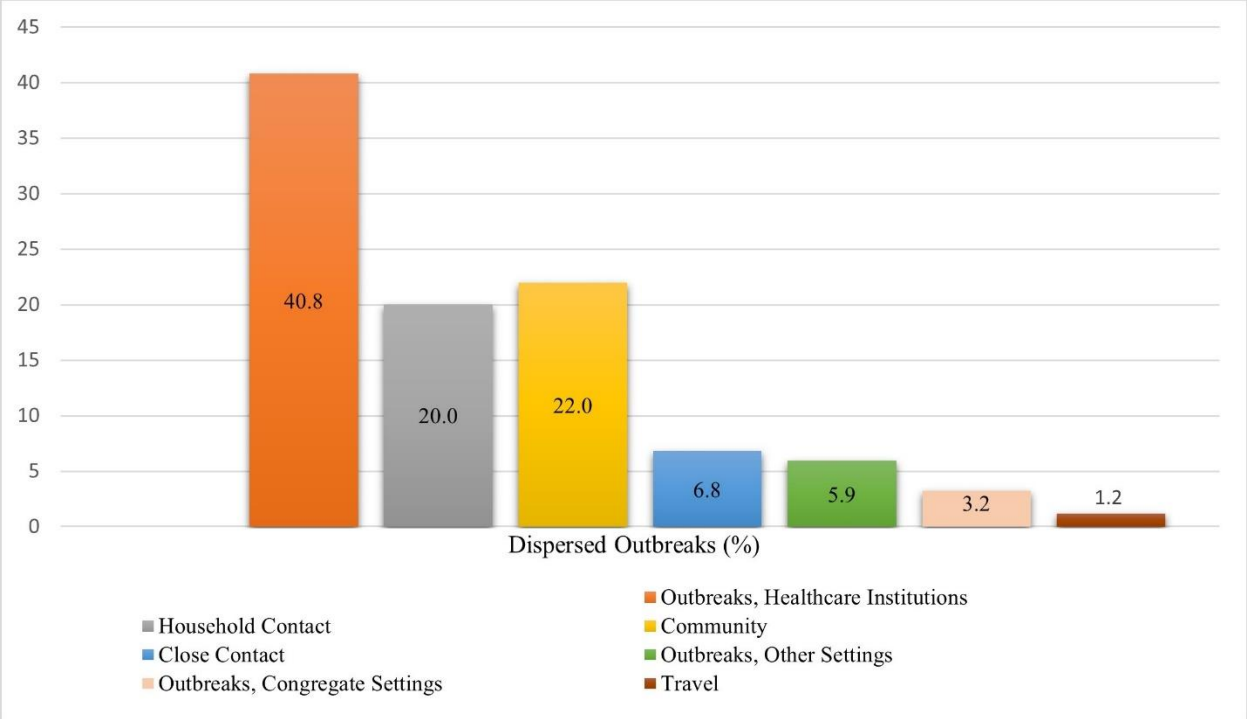
The results of the sensitivity analyses using the outbreak definition as the observed number of cases greater than the median number of cases show a similar spatial spread of the disease with a few exceptions (Fig S5.4). However, unlike the main analysis that used the mean number of cases in the outbreak definition, all neighbourhoods experienced an outbreak event using the median as the definition of an outbreak in the sensitivity analysis, since the median is only 9 cases, whereby the mean is 19 cases, much higher than the median number of cases. While reducing the temporal period from two weeks (61 dispersed neighbourhoods) to one week, we observed a lower number (31 neighbourhoods) of dispersed neighbourhoods (Fig S5.5).

5.4.2.2. Sources of Infection

To identify the sources of infection in the dispersed neighbourhoods, we included cases only for the 1st outbreaks in the dispersed neighbourhoods. A total of 26% of the cases did not have the source of infection in those neighbourhoods in the week of the first outbreaks, thus excluded from the analysis. About 40% of the source of infections in the dispersed neighbourhood was attributed to the healthcare institution (Figure 5.4). The majority of the sources of infection were households, close contacts and communities (49%) and after excluding healthcare institution-related outbreaks) it became 82%. Travel (1.4%), congregate settings (i.e., shelters, correctional

facilities) (3.9%), and other settings (i.e., workplace, daycare) (5.4 %) were not the major sources of infection in the study area.

Figure 5.4. Sources of infection of COVID-19 in the dispersed neighbourhoods in Toronto between January 2020 and June 2021.



.

5.2.3. Spatiotemporal Spreads during the First Three COVID-19 Waves

Figure S5.2 displays the spatiotemporal spread during the first three COVID-19 waves of the pandemic. In the first wave period, 58 (41%) index neighbourhoods and 55 (39%) dispersed neighbourhoods were observed. 23 (42%) of these dispersed neighbourhoods also became an index neighbourhood and dispersed to surrounding neighbourhoods. Similarly, in the second wave period, 54 (39%) index neighbourhoods and 52 (37%) dispersed neighbourhoods were observed. 22 (42%) of these dispersed neighbourhoods also became an index neighbourhood and

dispersed to the surrounding neighbourhoods. Out of these three waves, wave 3 experienced the maximum number of dispersion events throughout Toronto. During this third wave, 81 (58%) index neighbourhoods and 67 (48%) dispersed neighbourhoods were observed. 45 (67%) of these dispersed neighbourhoods also became an index neighbourhood and dispersed to surrounding neighbourhoods.

5.4.4. Classification of the Hotspots

The incidence rate and disease persistence of COVID-19 varied widely in the study area during our study period (Fig 5.5). The northwestern part of Toronto had a higher incidence rate with a long persistence. The southern part of Toronto had a relatively lower incidence rate, particularly in the neighbourhood no. 77 (Waterfront Communities – the island) but experienced a longer persistence of COVID-19. The spatial patterns of the incidence rate show that majority of the neighbourhoods had an overall incidence rate between 20-90 cases/1000 persons. Only a few neighbourhoods had an incidence rate of over 100 cases/1000 persons (Fig S5.1).

The hotspot classification chart displays how many neighbourhoods fall in each priority category based on epidemiological criteria used to classify the hotspots (Figure 5.6b). A total of 21 (15%) neighbourhoods were classified as high-priority (P1) hotspots. These neighbourhoods had a higher incidence rate and a higher persistence and were observed in the northwestern and northeastern corners of Toronto (Figure 5.6a). According to the 2016 census, a total of 530,119 population lived in the P1 hotspots. In total, 16 (11.4%) were classified as medium priority (P2/P3) hotspots. These neighbourhoods had either higher incidence rates with a shorter persistence or lower incidence rates with a longer persistence. P2/P3 hotspots were observed in different parts of Toronto. A large number of neighbourhoods [82 (58%)] were classified as low priority (P4)

hotspots because the neighbourhoods experienced a lower incidence rate with a shorter persistence.

Most of these hotspots were in the central and southwestern parts of Toronto.

Figure 5.5. Spatial distribution of incidence rate and persistence of the COVID-19 in Toronto, January 2020 – June 2021.

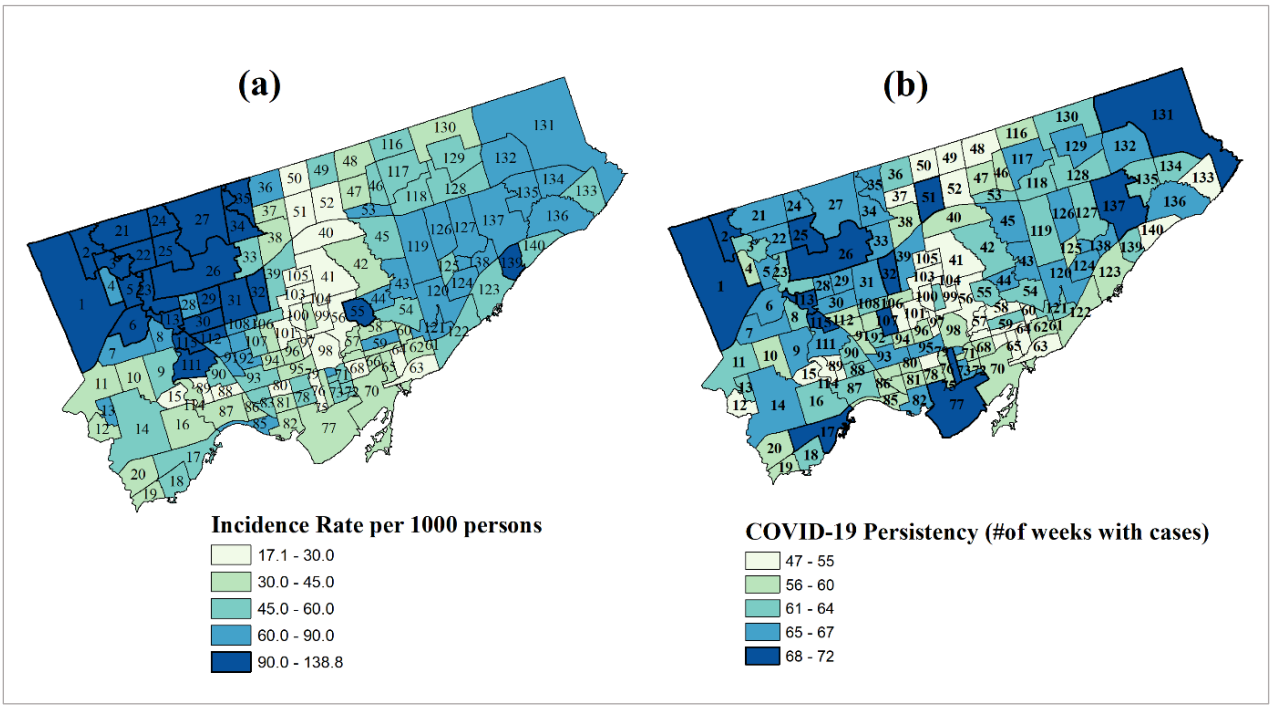
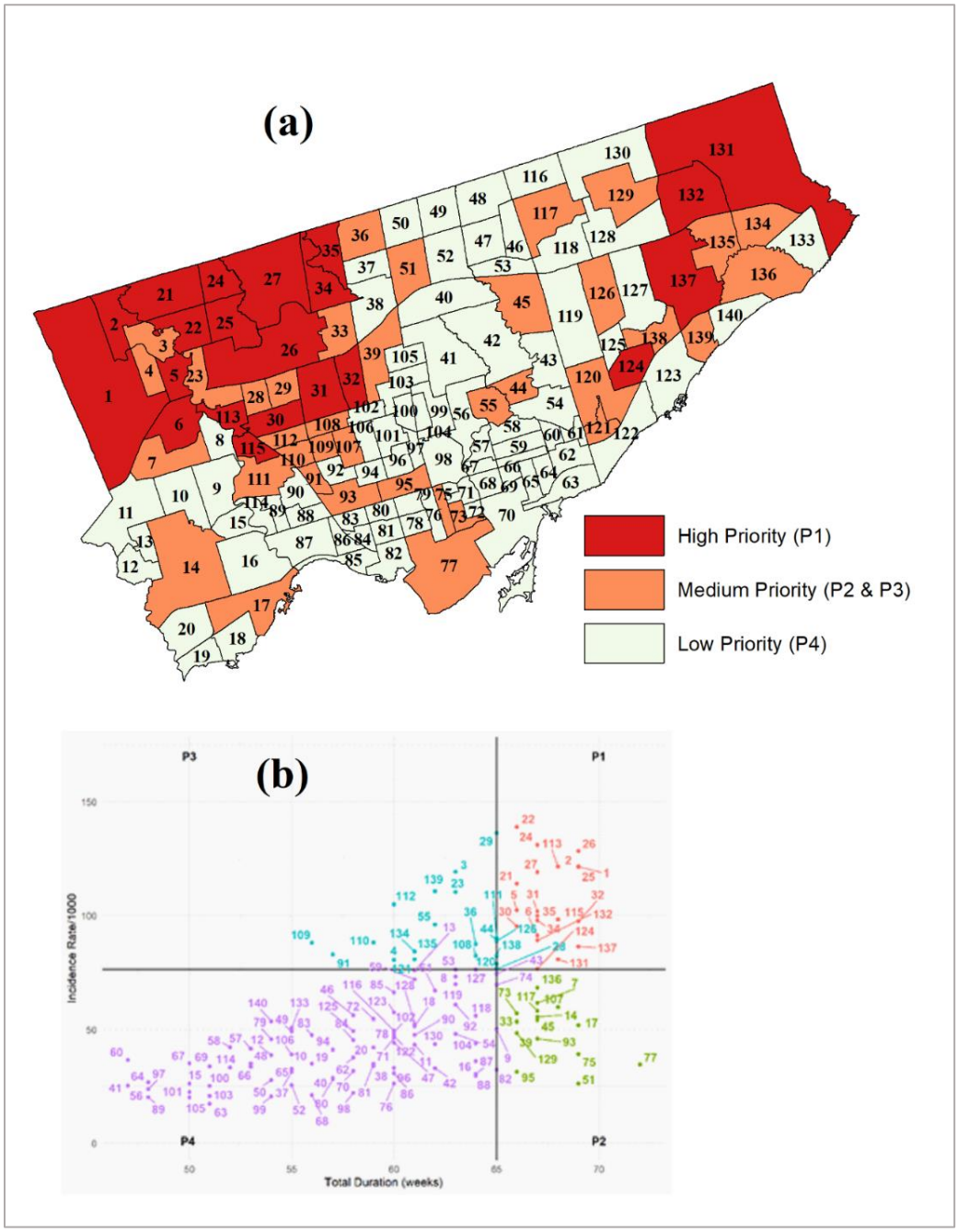


Figure 5.6. Hotspots (a) and classification chart (b) of COVID-19 incidence in Toronto between January 2020 and June 2021.



Note: The numbers in the chart and inside the map are the neighbourhood IDs.

5.4.4.1. Sensitivity Analyses of the Hotspots

The results of the sensitivity analyses using different cut-offs of the epidemiological parameters

show mostly similar results in identifying and classifying the hotspots, particularly in northwestern and northeastern Toronto (Figure S5.6). However, in contrast to the main analysis, a relatively higher number of P1 hotspots were observed mostly in the western and eastern parts, and a lower number of P4 hotspots were observed in the central part of Toronto.

5.5. Discussion

The model for the spatiotemporal spread of the disease developed in this study helped us visualize the spatial and temporal dynamics of COVID-19 in a densely populated urban city and provided a deeper insight into the spread of the COVID-19 outbreaks in the neighbourhoods of the city. This geographic model can be applied to any infectious disease to observe community-level spread patterns by updating the local definition of an outbreak and the incubation period of the disease for a geographic setting. This model can retrospectively observe the movement of outbreaks across neighbourhoods in space and time to identify possible sources and directions of the outbreaks. Understanding the spatial dynamics of the disease may help assist in planning timely interventions. Note that control, mitigation and eventually elimination of a highly transmissible disease such as COVID-19 require careful and comprehensive planning during the initial stages of the epidemic so that timely and efficient intervention plans can be executed. Our study suggests that the spread of the disease was not spatially random. Therefore, it is important to identify the areas where the targeted inventions are essential. This kind of model helps us identify the places where control efforts can also be initiated under resource constraint situations.

We observed that spatial dynamics of the COVID-19 outbreaks varied across the neighbourhoods in Toronto. An index neighbourhood possibly dispersed the outbreaks to a maximum of four neighbourhoods within two weeks. Some dispersed neighbourhoods became

index neighbourhoods and further dispersed the outbreaks to their nearby neighbourhoods. For example, in the south of Toronto in the Waterfront communities-the island (#77), the outbreak dispersed to five surrounding neighbourhoods within two weeks during the early stages of the pandemic and further continued to spread. Waterfront communities-island has a high population density (8943 persons per sq km), with a majority (69%) of the population at working age (24-52 years). This neighbourhood became an index neighbourhood and experienced dispersion to surrounding neighbourhoods during all three COVID-19 waves. The same neighbourhood also demonstrated a high disease persistence. In the northwestern part of Toronto, in the West Humber-Clairville neighbourhood bordering the Peel region, the outbreak spread to multiple neighbourhoods during the first wave of the pandemic. This neighbourhood with a relatively larger spatial size has a lower population density (1,117 persons per sq km) but a high population count of 33,112 with 82% of the population being a visible minority. Some similar spatiotemporal spread patterns were observed across the three COVID-19 waves, primarily in the southern, eastern and northern areas in Toronto where multiple dispersion events are observed. While the first two waves showed similar number of index and dispersed neighbourhoods, the third wave starting from March 5, 2021, was driven by the more transmissible Alpha (B. 1.1. 7) and showed a much higher dispersion events throughout Toronto.

The results suggest that ecological phenomena for creating a risk for the population are not random in space. Thus, understanding local-level variation is important for an effective control program. In our study, the first COVID-19 outbreak dispersion occurred between March 22 and March 28, 2020, in the Morningside neighbourhood (neighbourhood#135) and then spread to the Rouge neighbourhood (neighbourhood#131), where it became an index

neighbourhood of another dispersion event. The Morningside neighbourhood, located in the eastern part of Toronto, has a relatively low population density (3041 persons per sq km) and 73% of the population has a lower level of education. The northwestern and southern parts of Toronto experienced multiple outbreak dispersions during the early stages of the pandemic. Glenfield-Jane Heights neighbourhood (neighbourhood#25) in northwestern Toronto experienced the spread of the outbreak to four neighbourhoods between April 5 and April 11, 2020. This neighbourhood has a high population density with over 5,864 persons per sq km, with 76.6% of the population being a visible minority and 89% having a lower level of education. These findings may imply that early intervention at the index and its nearby neighbourhoods may have controlled the magnitude of the spread of the disease in the study area.

In our study, we attempted to provide an effective model for visualizing the spatiotemporal dynamics of the outbreaks, which is different from earlier studies that visualized infectious disease outbreaks. We believe that our model improves the traditional way of visualizing disease outbreaks using choropleth mapping or other methods such as SDE by including temporal structure, identifying index and dispersed neighbourhoods, dispersion directions and magnitude of the outbreaks in a single map. This kind of map may also be useful for evaluating the impact of an intervention, showing where intervention efforts are successful and where they are not. This could help to identify the barriers to a successful intervention program.

By analyzing the sources of infections of the COVID-19 outbreaks in the dispersed neighbourhoods, we identified a few factors that played a key role at a spatial scale in the spread of COVID-19 in Toronto. Based on the assessment of the sources of infection information from

the COVID-19 dataset, we observed that apart from outbreaks from healthcare institutions, most of the sources of infection in the dispersed neighbourhoods were from close contact either through individuals or from communities. Travel or workplace outbreaks had very little to contribute to the spread of the disease in the dispersed neighbourhoods in our study area. This suggests that close attention to the index and its nearby neighbourhoods could have controlled the spread of the disease in our study area.

When an outbreak hits a large area, it may not be practical to initiate an intervention in all the areas simultaneously. For instance, several efforts have been adopted by the Toronto local authorities to control the pandemic since the cases were reported at the Sunnybrook Health Sciences Centre in Toronto on January 23, 2020 (Silverstein et al., 2020). According to BBC News, on May 24, 2021, Toronto is considered to have the longest continuous COVID-19 lockdown among any major city around the world ('Toronto Lockdown - One of the World's Longest?', 2021). On March 23, 2020, the Toronto Mayor declared a citywide emergency, restricting businesses from operating. Because of the spread of the disease, more regionally targeted restrictions and policies were applied. These measures included lockdowns, mask mandates, closure of schools, indoor gathering limits, limited non-essential business, limiting shopping malls to curbside pickup, and preventing any subsequent large gatherings (J. A. Long et al., 2021; Soucy et al., 2021). The city of Toronto prioritizes the postal code areas with high neighbourhood-level case rates for mobile and pop-up clinic vaccinations (City of Toronto, 2021b, 2022b). Given all those efforts, there were no signs of declining the disease in Toronto. Past studies have found that spatially targeted interventions can be highly effective for controlling infectious disease outbreaks (Cudahy et al., 2019; Finger et al., 2018; Franch-Pardo

et al., 2020; Khundi et al., 2021). The results of our hotspot analysis provided important data by which one could prioritize areas for local interventions.

Along with the incidence rate, our approach includes the persistence of the cases during the study period in defining the hotspots. In our study, the high-priority areas were observed near the western part of Toronto, bordering the Peel region, which had reported a high COVID-19 incidence rate in Canada. It is possible that Toronto was affected by the high number of cases in the Peel region. The central regions of Toronto near downtown observed a few dispersions but were mostly at low risk, whereas a higher risk was observed near the outer boundaries.

In an urban area, various risk factors such as socioeconomic inequalities, demographic, environmental factors or pre-existing health conditions in a population are often used to explain the variations in risk and outbreak dynamics (Ingen et al., 2021; News ·, 2020; Vaz, 2021). Population density is an important factor in the spread of COVID-19 incidences (Bhadra et al., 2021; Kadi & Khelfaoui, 2020; Pequeno et al., 2020; Sy et al., 2021). We observed that many of the spread of outbreaks occurred in the highly densely populated areas (Central and Southern Toronto) with a population density ranging from 6500 – 44321 persons per sq km. However, in the northeastern and northwestern corners of Toronto, where the population density is lower (1000-3000 persons per sq km) than in the other part of Toronto, we observed the high-risk priority neighbourhoods and multiple dispersions, suggesting that population density alone may not explain the increased risk of COVID-19. Targeting the high-priority hotspots first, as defined in this study, and subsequently to medium-priority hotspots and then the low-priority hotspots would be the most efficient strategy in controlling the disease.

This study is not free from limitations. We used the commonly used definition for the outbreak. The recommendations of international policy guidelines and research agendas are based on the perceived standardized definition of an outbreak characterized by a prolonged, high-caseload, extra-seasonal surge. Any changes in the definition may change the dynamic patterns to some degree. However, the results of our sensitivity analysis showed only minor variations in the spatiotemporal dynamics of the outbreak. Second, asymptomatic cases are not included in the analysis, as these data were unavailable for our study area.

Additionally, access to testing and testing resource allocation across neighbourhoods may have varied over time (CBC News, 2020; Ontario, 2022). These missing numbers of cases could potentially result in an underestimation of disease transmission. Finally, we did not have the individual or neighbourhood-level mobility data to explain some of the disease spread processes. However, we have seen travel contributed only a little in explaining the spread of the outbreak.

5.6. Conclusion

In a wider outbreak control effort, research agendas and subsequent policy guidelines have heavily focused on methods to identify and predict outbreaks and how to respond appropriately by outbreak response protocols to better plan for future outbreak occurrences using effective control measures. With appropriate and timely control, the spread of the disease outbreak can be minimized. Our study offers a new approach to a deeper understanding of the spatial dynamics of infectious disease outbreaks, which would facilitate the intervention program for controlling the spread of the outbreaks. The approach could be useful for understanding spatial dynamics of other infectious diseases where local level spatially enabled data are available.

Chapter 6 Discussion and Conclusion

6.1. Discussion

6.1.1. Key Findings

This dissertation is composed of three original research manuscripts exploring the spatial epidemiology of COVID-19 outbreaks, the urban environment and public health planning using spatial and spatiotemporal methods. The manuscripts in Chapters 3 to 5 significantly contribute to the common theme, neighbourhood-level assessment of the spatial and spatiotemporal epidemiology of COVID-19. The key findings and major contributions for each of the three chapters are briefly summarized in Table 6.1, with reference to the original research objectives and questions provided in Table 1.1 in Section 1. More details are given in the following sections of this chapter.

Table 6.1. Key findings and major contributions of the thesis.

Chapter #	Key Findings	Major Contributions
Chapter 3 (Manuscript 1)	1. Risks of COVID-19 vary in space-time. 2. At the global scale, lower levels of education and the percentage of immigrants were found to be positively associated with COVID-19 transmission rates. 3. At the local scale, we found variations in the association between COVID-19 and various socioeconomic risk factors.	1. Provides an understanding of COVID-19 disease epidemiology in Toronto. In particular, where are the high and low risk areas located in Toronto, how socioeconomic inequalities in an urban environment can influence disease transmissions and how these associations can vary locally. 2. Provides the methodological framework by applying a local multiscale regression modelling approach incorporated with a correction method to identify the local variations in the association between COVID-19 and socioeconomic risk factors in an urban environment. 3. Provides insights into hotspots and important risk factors for strategies and interventions for public health planning.
Chapter 4 (Manuscript 2)	1. LST is found to be positively associated with COVID-19.	1. Provides empirical evidence of the association between urban land surface temperature and COVID-19.

	2. A lower level of education is positively associated with COVID-19. 3. High-risk areas or hotspots were identified. 4. Higher temporal risk was observed during Spring 2021. 5. The spatiotemporal risk patterns identified that 60% of neighbourhoods had a stable, 37% had an increasing, and 2% had a decreasing trend over the study period.	2. Provides an understanding of COVID-19 disease epidemiology in Toronto. In particular, where are the high and low risk areas located in Toronto, how socioeconomic inequalities in an urban environment can influence disease transmissions. 3. Provides a methodological framework by applying Bayesian infectious disease modelling technique and remote sensing technology for deriving temporal, spatiotemporal trends and area-specific relative risk. 4. Provides insights into hotspots and risk factors for strategies and interventions for public health planning.
Chapter 5 (Manuscript 3)	1. Visualizes the spatial spread of COVID-19 from index neighbourhoods to dispersed neighbourhoods within two weeks. 2. Created high, medium and low-priority hotspots for COVID-19 using disease persistence and incidence rate.	1. Provides an understanding of the community-level spatiotemporal spread patterns of an infectious disease in Toronto, particularly how these spread dynamics can be heterogeneous in space and time. 2. Applies a novel spatiotemporal model to visualize the outbreak spread patterns. 3. Provides insights into the index or origin neighbourhoods and priority-based hotspots for interventions and strategies for public health planning.

The first manuscript, titled "*Spatial and spatiotemporal clusters and the socioeconomic determinants of COVID-19 in Toronto neighbourhoods, Canada: An Observational Study*," examined if COVID-19 risks varied across space and time, identified the global socioeconomic risk factors of COVID-19 and inquired how the associations between COVID-19 transmissions and socioeconomic variables varied across different neighbourhoods. This study uses test-based and model-based methods to identify the spatial patterns and risk factors in Toronto. The multiscale geographically weighted regression with a corrected method identified these local

variations in the association between the variables. This study found lower education and concentration of immigrants to be significant contributors to COVID-19. Additionally, even though income and unemployment were not found to be significant contributors at the global scale, the local regression model found these risk factors to be significant contributors in a few neighbourhoods. The study identified several high-risk space-time clusters in the northwest and eastern part of Toronto.

The second manuscript, titled "*Identifying Spatiotemporal Patterns of COVID-19 and the Drivers of the Patterns in Toronto: A Bayesian Hierarchical Spatiotemporal Modelling*", quantified the area-specific risks and trends by using satellite-driven time series LST data and socioeconomic variables to predict COVID-19 risk in a small area. The best-fitting spatiotemporal model used spatial, temporal and type II space-time interaction effects to show the temporal and spatiotemporal trends of COVID-19 risk. The study found LST to be positively associated with COVID-19. Similar to manuscript 1 (Chapter 3), education was also found to be an important contributor to COVID-19 risk. The spatial patterns of risk were heterogeneous in space with multiple high-risk neighbourhoods in Western and Southern Toronto. The temporal trend analysis identified high risk was observed during Spring 2021. The spatiotemporal risk patterns identified that 60% of neighbourhoods had a stable, 37% had an increasing, and 2% had a decreasing trend over the study period. Most neighbourhoods experienced either stable or increasing spatiotemporal risk during the study period.

The third manuscript, titled "*Modelling the spatiotemporal spread of COVID-19 outbreaks and prioritization of the risk areas in Toronto, Canada*", mapped the spatiotemporal spread of COVID-19 spread at the neighbourhood level using a novel modelling technique and created a

priority scheme for the hotspots by using incidence rate and disease persistence. This study observed several spatial spreads of the disease in Toronto within two weeks of the onset of the first outbreaks from the index neighbourhood to its first-order neighbourhood networks. Apart from the healthcare institutions, most of the sources of infection were from the households and communities (82%) in the dispersed neighbourhoods, suggesting an expansion of transmission from the nearby neighbourhoods.

6.1.2. Comparison of the findings

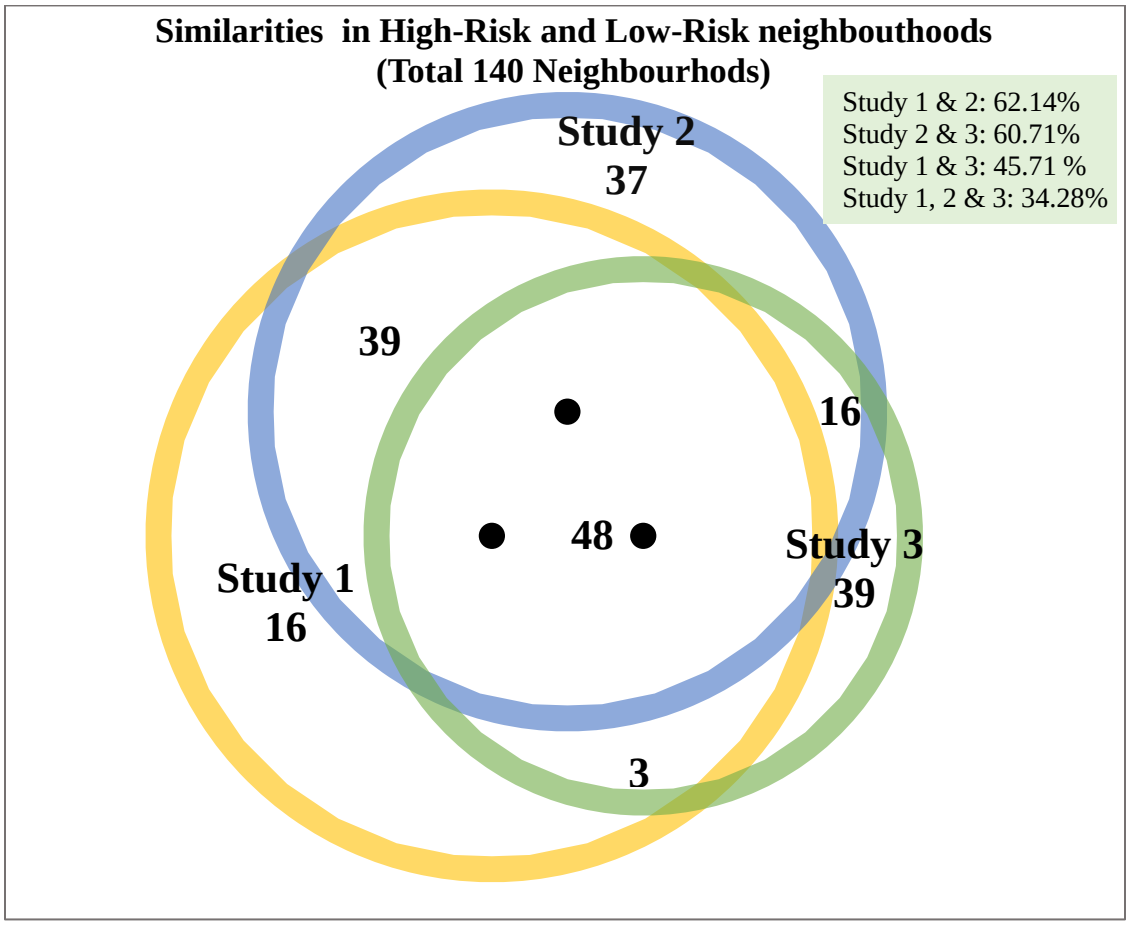
The three research studies conducted in this dissertation use three different methods to identify the COVID-19 spatial risk in Toronto. I have further analyzed the similarities and differences between the spatial risk of COVID-19 in the neighbourhoods among these three studies (Figure 6.1, Table 6.2). Cohen's Kappa Coefficient test was also performed to assess the degree of concordance (similarities) or discordance (differences) between the risk areas identified in each study.

Study 1 in Chapter 3 adopted the SaTScan, a widely popular frequentist scan statistic method, to identify the space-time risk areas. From the outputs of this study, neighbourhoods with a significant relative risk were classified as high-risk neighbourhoods and neighbourhoods without a significant relative risk were classified as low-risk neighbourhoods for analysis. Study 2 in Chapter 4 adopted the Bayesian hierarchical modelling technique while adjusting for socioeconomic and LST variables to yield the area-specific relative risk. From the outputs of this study, the neighbourhoods with an area-specific spatial relative risk >1 were classified as high-risk neighbourhoods, and the remaining neighbourhoods were classified as low-risk neighbourhoods for analysis. Study 3 in Chapter 5 adopted a hotspot prioritization method, including incidence rate and disease persistence, to calculate the high, medium and low-priority

risk neighbourhoods. From the outputs of this study, the neighbourhoods with high or medium-priority risks were classified as high-risk neighbourhoods, and neighbourhoods with low-priority risks were classified as low-risk neighbourhoods for analysis. These neighbourhood-specific spatial risks of COVID-19 were mapped in Sections 3.4.2, 4.4.3 and 5.4.3.

Figure 6.1 uses a 3-circle Venn diagram to show the similarities in high and low-risk neighbourhoods between (i) study 1 and study 2, (ii) study 2 and study 3, (iii) study 3 and study 4 and (iv) study 1, 2 and 3. The numbers inside the Venn diagrams denote the common number of neighbourhoods with a high/low risk between studies. Overall, 48 (34.2%) neighbourhoods showed similar geographic patterns of risk across all three studies. These similarities can be observed primarily in the northwestern and northeastern parts of Toronto. Cohen's Kappa coefficients were calculated to evaluate the degree of concordance between the geographic patterns of risk neighbourhoods across the three studies (Cohen, 1960). The Kappa coefficient value evaluates the degree of concordance (similarities) or agreement between the neighbourhood's risk patterns across the studies.

Figure 6.1. Similarities in the High and Low-Risk Neighbourhoods across the three studies.



Notes: Study 1: SaTScan, Study 2: Bayesian Hierarchical Modeling, Study 3: Hotspot Prioritization.
The numbers inside the Venn diagram denote the common number of neighbourhoods with a high/low risk between studies.

Table 6.2. Agreement and disagreement in the High/Low-Risk Neighbourhoods among the three studies. (Study 1: SaTScan, Study 2: Bayesian Hierarchical Modeling, Study 3: Hotspot Prioritization).

	Concordance and Discordance pairs				Agreement (%)	Kappa (95% CI)
	high/high	low/low	high/low	low/high		
Study 1 & 2	56	31	42	11	62.14%	0.25 (0.11, 0.39)
Study 2 & 3	35	50	32	23	60.71%	0.20 (0.04, 0.37)
Study 1 & 3	40	24	58	18	45.71%	-0.01 (-0.15, 0.12)
Study 1, 2 & 3	27	21				

Notes: According to the values of the Kappa coefficient, which varies between 0 and 1, the degree of concordance will be determined as follows: poor ($k < 0.2$), fair ($0.2 \leq k < 0.4$), moderate ($0.4 \leq k < 0.6$), good ($0.6 \leq k < 0.8$); excellent ($k > 0.8$), and perfect ($k=1$).

Cohen's Kappa Coefficient test result suggests that study 1 and study 2 had 62.14% similarities in the risk patterns with a Kappa value of 0.25 (0.11,0.39), suggesting a fair concordance. These two studies have common high-risk areas mostly in the northwestern and northeastern neighbourhoods of Toronto and common low-risk areas in northern and southern neighbourhoods of Toronto. Study 2 and study 3 had 60.71% similarities in the risk patterns, with a Kappa value of 0.20 (0.04, 0.37), suggesting a fair concordance. These two studies have common high-risk areas mostly in the northwestern, southwestern, northeastern and southern (neighbourhood #77) neighbourhoods of Toronto and common low-risk areas in central and southern neighbourhoods of Toronto. Finally, study 1 and study 3 had 45.71% similarities in the risk patterns, with a Kappa value of -0.01 (-0.15, 0.12), suggesting a poor concordance. These two studies have common high-risk areas mostly in the northwestern and northeastern neighbourhoods of Toronto and common low-risk areas in the central neighbourhoods of Toronto.

The results of the hotspot comparison assessment suggest that SaTScan produces relatively more similar risk patterns (62.14%) to the Bayesian modelling method compared to the hotspot prioritization method (60.71%). The outputs of the SaTScan and the hotspot prioritization methods showed a higher level of discordance (54.29% dissimilar). Various differences were also observed among the high or low risk areas across the three studies. For example, in the southern part of Toronto, in Neighbourhood #77, a very high spatial risk is observed in Study 2, and a medium priority risk was observed in Study 3. However, in Study 1, using the SaTScan method, this neighbourhood did not have a significant risk in this

neighbourhood. Such a difference is also expected, as the mathematics behind identifying the clusters or risk calculation is not the same in the applied methods.

A major limitation of the SaTScan method is that SaTScan can not detect non-circular or irregular shaped clusters and, therefore, requires a circular pattern of the cluster with the likelihood ratio statistic calculated from the number of observed and expected cases from within and outside the clusters (J. Chen et al., 2008). It is possible that the Bayesian and the SaTScan method have some similar performances in identifying circular hotspot clusters that can be noticed in the northwestern and northeastern parts of Toronto. Bayesian methods are expected to have more flexibility regarding non-circular hotspots (Lawson, 2010). The methods also differ in terms of using the geographic locations of each neighbourhood and their neighbouring neighbourhoods. The Bayesian method basically smooths the data compared to the SaTScan and the hotspot prioritization methods (Lawson, 2010). The correlated component of the Bayesian BYM model depends on the adjacent neighbourhoods regardless of the distance and population and is, therefore a bit more flexible than SaTScan (Aamodt et al., 2006). At the same time, the hotspot prioritization method does not factor adjacent areas into account while calculating risk.

In contrast to the Bayesian BYM model, the circular scan in the SaTScan method does not take into account whether the neighbourhoods within the circles are neighbourhoods (Aamodt et al., 2006). Therefore, it is possible that SaTScan may detect relatively large clusters of high-risk areas compared to the Bayesian method and the hotspot prioritization method which was observed in this dissertation. Furthermore, in contrast to the SaTScan and the hotspot prioritization method, the Bayesian method adds correlated spatial random effects into the

model, which can handle overdispersion or extra spatial variability in the data (Gschlößl & Czado, 2008).

There are also differences among these methods in terms of practicalities. The SaTScan and hotspot prioritization methods are relatively simple to implement and require relatively less computational time than Bayesian methods (Aamodt et al., 2006). Additionally, unlike the Bayesian method, the SaTScan and the hotspot prioritization method do not require prior, which is often unknown for a new disease (Best et al., 2005). The SaTScan has the ability to uncover spatial and temporal clusters using various distribution models its ability to uncover spatial and temporal clusters using various distribution models (Coleman et al., 2009). SaTScan is geared to identify space-time clusters, whereas the Bayesian method has space-time structures built into it but is targeted to identify area-specific risk (Best et al., 2005; Hohl et al., 2020). Specifying the BYM model in the Bayesian method also requires finding the adjacent neighbourhoods for all neighbourhoods in the study area (Aamodt et al., 2006). The choice of the appropriate method to identify high-risk areas in a study may depend on various factors. Such as: (i) computational capacity, (ii) time, (iii) assumed shape of the risk clusters (iv) data quality and limitations, or (v) methodological strengths and limitations.

For our case study, Bayesian methods seem more appropriate for identifying the hotspots. In a small-area analysis, data can be sparse, and a small population can have a smaller number of observed and expected cases, creating unstable risk estimates (Beale et al., 2008). In the study area in Toronto, the shape, size and population of the 140 neighbourhoods in Toronto vary. Also, a limitation of COVID-19 datasets is the absence of asymptomatic or unreported cases, which could also be responsible for significant transmission (Alene et al., 2021). The Bayesian

approach allows accounting for data sampling variability, parameter uncertainty, and potential dependence structures such as spatial, temporal and spatiotemporal structures. In Bayesian modelling, it is easier to incorporate variance components hierarchically and provide better estimates of predictive uncertainties than the other methods discussed in this dissertation. Therefore, a Bayesian method used in this study might have been a more reasonable option for modelling the neighbourhood scale risk of COVID-19 in Toronto or other similar settings.

In Chapter 3, the global relationship between COVID-19 and socioeconomic factor were identified using the widely used spatial regression models (SEM and SLM), while in Chapter 4, the same relationships were assessed using a Poisson-based Bayesian Hierarchical model. Both studies have identified lower levels of education to have a positive association with COVID-19 incidence. However, in Chapter 3, the non-Bayesian spatial model identified percentages of immigrants to be positively associated with COVID-19, whereas no associations were found from the Bayesian modelling method in Chapter 4.

It is important to note that the socioeconomic variables were selected based on the literature review and availability of such data, specifically during the study period and at the neighbourhood level. Unfortunately, there could be some other important covariates which were not adjusted into the models. The robust Bayesian spatiotemporal hierarchical models used in Chapter 4 with structured, unstructured, random effects allowed compensating for the missing/unobserved covariates in the models, which may not be the case for the spatial regression model in GeoDa (Knorr-Held, 2000). In particular, the Type II space-time interaction in the final Bayesian model in Chapter 4 implied that the missing covariates have smoothly varying structures through time but are highly localized in their effect on the outcome and so

have no structure over space (Haining & Li, 2020). Therefore, it is possible that not compensating for the missing covariates may have caused a false positive association in the GeoDa model for the percentage of immigrants in terms of overall risk. Therefore, a Bayesian spatiotemporal model with various spatial, temporal and spatiotemporal effects using space-time interaction terms might be a more plausible option to model the global associated risk factors that compensate for potential missing covariates.

6.1.3. Residual Assessment of the MGWR Model

In Chapter 3, the log of the COVID-19 incidence rate was used in the M(GWR) model to satisfy the assumption of linearity to fit the data into the model. Using the rates is crucial in ecological studies as it controls for the size of the individual objects since using absolute values without adjusting them with the population at risk, for instance, may lead to unrealistic correlations (Tiefelsdorf & Kyunghee Rhyu, 2022; D. Wheeler & Tiefelsdorf, 2005). However, the varying denominators of the rates, generally the population at risk of a geographical unit, can also make these rates inherently heteroscedastic (Tiefelsdorf & Kyunghee Rhyu, 2022). To control for the varying variances of the observations, the M(GWR) model standardized the data by performing a z-transformation on the dependent and the local independent variables so that each variable has a mean of 0 and a standard deviation of 1 (T. M. Oshan et al., 2019). Additionally, a basic assumption of the M(GWR) models is that the error terms of the models are independent and normally distributed (Fotheringham et al., 2017; H. Yu et al., 2020). The presence of spatial autocorrelation of the error terms can therefore invalidate these statistical assumptions of homoscedasticity of the error in an M(GWR) model and, ultimately, mislead the results of the statistical performance (Brunsdont et al., 1998; Leung et al., 2003; Mei, 2004). To avoid this

issue, testing for spatial autocorrelation among the residuals of a geographically weighted regression is crucial (Mei, 2004). In this dissertation, a widely used Global Moran's *I* test (F. Dormann et al., 2007; Leung et al., 2003; Mei, 2004) was used to further explore the spatial autocorrelation of the residuals obtained from the M(GWR) model (Chapter 3), and the results are presented in Appendix C. These results imply that the assumption of the homoscedasticity of the error residuals of the M(GWR) model was not violated.

Overall, the M(GWR) is a considerable improvement over the previous traditional GWR methods and comparable to the Bayesian spatially varying coefficient model (Finley, 2011; T. Oshan et al., 2019; H. Yu et al., 2020). A limitation of the M(GWR) model is that it does not obtain the parameter estimates that project the observed response vector into the predicted response vector (H. Yu et al., 2020). Another limitation of the M(GWR) is that it may be difficult to refine the scale of action from a large sample, particularly when it takes a data set larger than 10,000 observations.(Fotheringham et al., 2017; H. Yu et al., 2020). Furthermore, geographically weighted regression models are more prone to multicollinearity. In the M(GWR) model, multicollinearity can occur in different situations, such as one of the explanatory variables is spatially clustered or two or more explanatory variables are highly correlated (H. Yu et al., 2020). However, M(GWR) yields lower AIC and AICc values compared to the traditional GWR and is less prone to multicollinearity issues, but the processing time is much longer for M(GWR) compared to GWR, and it increases with the size of the data (Fotheringham et al., 2017; T. M. Oshan et al., 2020).

6.2. Contributions

As summarized in Table 6.1, this dissertation provided significant contributions, which are detailed below. From a spatial epidemiological perspective, this dissertation sought to advance the understanding of the relationship between a highly transmissible disease and the urban socio-environment. From an analytical perspective, this dissertation was motivated to apply advanced modelling techniques for analyzing COVID-19 and urban neighbourhood scale data. From a public health policy perspective, this dissertation looked to provide information that could help design and implement urban public health intervention policies and programs by the local public health. The contributions of this dissertation against each objective are discussed in detail below.

6.2.1. Epidemiological contributions

Overall, this dissertation contributed to the theories in the field of infectious disease research, and public health sciences focused on the relationship between the urban environment and COVID-19 in three ways. First, this work measured how COVID-19 disease risk and trends varied in space and time in Toronto to advance past research. Second, this work mapped how outbreaks spread over space and time within neighbourhoods in Toronto. Third, this work identifies the important socioeconomic risk factors of COVID-19 in Toronto, providing evidence of neighbourhood effects and also asks whether variations in urban heat or land surface temperature in Toronto can be used to predict COVID-19 risk. Across the three manuscripts, this dissertation strengthens evidence from past neighbourhood-based research on how COVID-19 risks, patterns and trends are heterogeneous in space and over time in Toronto neighbourhoods (Kan et al., 2021; B. Li et al., 2021; Perles et al., 2021). Each of the three research manuscripts also makes some unique contributions.

The first (Chapter 3) and second (Chapter 4) manuscripts support historical and contemporary evidence of neighbourhood inequalities in a pandemic by suggesting that urban socioeconomic inequalities across neighbourhoods were associated with increased infectious disease risk (Bambra et al., 2020; Cordes & Castro, 2020, 2020). The first manuscript found that higher concentrations of foreign-born immigrants are positively associated with disease transmission rates in Toronto, which supports previous work that found that COVID-19 transmissions disproportionately impact immigrant communities (Clark et al., 2020; Hasan Bhuiyan et al., 2021; Oluyomi et al., 2021; Strully et al., 2021).

This study hypothesized that COVID-19 disproportionately impacted the foreign-born immigrant population, which could be attributed to immigrants living in overcrowded housing conditions, low-income neighbourhoods or working in environments where physical distance may be difficult (Choi et al., 2020; CTV, 2020). The first and second manuscripts both supports and strengthens past studies on the importance of the level of education in understanding and awareness of the risks of fighting a pandemic (Borjas, 2020; Fielding-Miller et al., 2020; Goutte et al., 2020). The dissertation hypothesizes that lower education levels may increase COVID-19 risk indirectly through health-seeking behaviour and efficiency in navigating the health systems. Furthermore, the first manuscript identifies the locally varying coefficients for each socioeconomic risk factor by identifying neighbourhood-specific associations. It was observed that the unemployment rate and poverty level were also important influencers of COVID-19 transmission in some neighbourhoods in Toronto. This builds on past studies that focused only on global risk and overlooked how different factors may impact disease transmission differently in each neighbourhood within a small area (Boland et al., 2021; Cao et al., 2020; Castro et al.,

2021; Sannigrahi, Pilla, Basu, & Basu, 2020; Sannigrahi, Pilla, Basu, Basu, et al., 2020b; Y. Sun et al., 2020b).

The second manuscript (Chapter 4) identified land surface temperature to have positive associations with COVID-19 transmission in Toronto. This finding supports a prior study by Hassan et al. (Hassan et al., 2021) that found a positive association while challenging a study that identified a negative association (Johnson et al., 2021). LSTs are also more spatially representative compared to ambient temperature collected from the ground weather stations from one location (do Nascimento et al., 2022), and the finding provides empirical evidence to establish the link between temperature and the COVID-19 surge in Toronto. This study also identifies that the temporal and spatiotemporal trends of disease risk were heterogeneous over space and time, mostly showing stable or increasing patterns in Toronto. The temporal risk fluctuation over time in the study area provided an understanding of how risks can vary in time in Toronto, highlighting the importance of incorporating temporal structures in future research. We identified significant changes in the risk of COVID-19 six times across the study period in Toronto with varying trends. This finding suggests that the temporal trend of the epidemic of COVID-19 is different from other coronavirus diseases, such as the SARS (severe acute respiratory syndrome) epidemic in 2003, in which case the number of reported SARS cases has increased exponentially over time, and the outbreak lasted a shorter period of approximately six months (Boulos, 2004; CDC, 2021; Galvani et al., 2003; X. Zhang et al., 2020). This research observed the spatiotemporal risk trends of COVID-19 (stable, increasing or decreasing), which was hardly observed in prior studies. The temporal risk was particularly high in early spring and

lowest in summer, thus advancing the understanding of the seasonal impact on disease transmission in Toronto.

In Chapter 5, the spread mechanism of COVID-19 is visualized in a single map that shows the spatial and temporal dynamics of COVID-19 in Toronto. The findings advanced the understanding of the spread dynamics, magnitude and directions of COVID-19 outbreaks in Toronto. The study provides theories for public health planning programs, highlighting the need for early intervention programs at the neighbourhood-level targeting sources areas to stop spreading into surrounding communities.

6.2.2. Analytical contributions

This dissertation has developed and applied advanced modelling tools that provide the framework for the analysis of spatial and temporal data of COVID-19 at the neighbourhood scale in Toronto. While the first manuscript (Chapter 3) adopts a multiscale geographically weighted regression modelling framework to model urban neighbourhood scale socioeconomic inequalities' impact on disease rates, the second manuscript (Chapter 4) applies the Bayesian hierarchical modelling approach using a set of observed covariates and using spatial, temporal and multiple space-time random effects (Knorr-Held, 2000; Knorr-Held & Besag, 1998). The third manuscript (Chapter 5) creates a novel visualization tool that extends existing work of disease mapping that generally compares multiple maps from different time points and identifies the spread dynamics in one map.

Chapter 3 finds that the MGWR model was the better-fitted model to explain the local variations in associations between socioeconomic variables and COVID-19 risk. This research overcomes the common limitations of the global regression models by taking the non-stationarity

of spatial associations between disease rates and explanatory variables into account (Thayn & Simanis, 2013). This study strengthens the traditional geographically weighted regression methods by applying the merits of MGWR, which has been shown to be a viable model form that generates different measurements of scale for different processes, is computationally less demanding and arguably intuitive (Fotheringham et al., 2017). The MGWR method further improves upon the majority of the studies that identified the local variations of COVID-19 by incorporating a correction method to incorporate the multiple testing issues (A. R. da Silva & Fotheringham, 2016).

Chapter 4 contributes to a robust methodological framework for fitting spatiotemporal LST data into a Bayesian model and provides a framework for understanding the temporal and spatiotemporal trends of COVID-19 risk. The presented Bayesian spatial and spatiotemporal statistical approaches have been widely applied in spatial epidemiological research but not necessarily in understanding spatiotemporal trends of COVID-19 (Franch-Pardo et al., 2020; Nazia et al., 2022). This study provided a flexible analytical framework for analyzing the neighbourhood COVID-19 risk trends to show if COVID-19 risk shows increasing, decreasing and stable patterns across neighbourhoods. In particular, this research strengthens and stabilizes estimation by proving a modelling approach for evaluations of spatiotemporal data of LST. This study improves upon the existing modelling of LST spatiotemporal data from previous COVID-19 research by acquiring satellite images using longer temporal data (79 consecutive weeks), higher spatial resolution (100 meters) and a smaller spatial scale (neighbourhood) and applying atmospheric corrections to control for errors in LST calculation (Das et al., 2020). We also used atmospheric corrections methods on these images by adopting Sobrino et al. in 2008's Land

Surface Emissivity (LSE) model (Sobrino et al., 2008), which provides a high estimation accuracy based on previous studies (Z.-L. Li et al., 2013; Sobrino et al., 2008; X. Yu et al., 2014; Zhao et al., 2009). This study contributes to the existing COVID-19 literature by providing a methodological framework using higher spatial and temporal resolution satellite imageries at a local spatial scale, which is more spatially representative and may have provided a more accurate estimate due to the use of the atmospherically corrected data on LST.

In Chapter 5, I developed a novel geographic model that can be applied to investigate neighbourhood or community-level spread patterns of infectious diseases by updating the local definition of an outbreak and the incubation period to adjust for local settings (Gianquintieri et al., 2020; R. C. Jones, 2021; Kang et al., 2020). Previous studies hardly modelled urban neighbourhood scale spread dynamics of COVID-19. Compared to prior studies that visually compared multiple maps, this study provides an analytical framework for neighbourhood-level visualization of the spatiotemporal spread of infectious diseases and identifies the outbreak source neighbourhoods (Cuadros et al., 2020; Y. Feng et al., 2020; Gao et al., 2021; H. Yu et al., 2021). The methodological framework applied here can also be used in other urban neighbourhood-level settings or for other infectious diseases. This research also created a prioritization scheme that includes the disease persistence parameter which has been used in the past for other infectious diseases but rarely used in COVID-19 research (Debes et al., 2021; Ngwa et al., 2021).

This study also explores different methodologies to identify high-risk areas and discusses the differences in the outputs in Section 6.2. The methodological frameworks and suggested

recommendations may be adopted by public health planners by integrating with local knowledge and computational capacities.

6.2.3. Public Health planning contributions

This dissertation provided evidence through epidemiological theories and analytical tools for urban public health planners to design and implement different intervention plans and programs in the Canadian urban context. Collectively, the three manuscripts provide four major contributions to public health planning. First, this research identified local and global risk factors that contributed to high risks in neighbourhoods. Second, temporal and spatiotemporal trends of COVID-19 risks are identified, which could inform health officials about planning efforts. Third, high-risk areas are identified where policies and more programs are necessary to decrease risks in those areas through different plans and programs. Four, the neighbourhood-wide spread dynamics of COVID-19 outbreaks can inform public health officials to target the source neighbourhoods.

Focusing on the public health policy contributions, in Chapter 3, the local associations between the four socioeconomic variables and COVID-19 are visualized in section 3.4.3.2. The findings can allow public health officials to plan whether a particular type of intervention might be required to address the neighbourhood's specific socioeconomic vulnerabilities. For instance, that low-income and unemployment rates were significant in a few neighbourhoods, which could account for public health planning in those particular neighbourhoods.

Chapters 3 and 4 identified education level as an important global determinant of the heightened risk of COVID-19; therefore, factoring education into planning programs could be an

important part of pandemic planning strategies in urban areas. Additionally, previous studies have linked higher education levels with vaccine hesitancy (Mathieu et al., 2021; Suthar et al., 2022; Watson et al., 2022). One way to address this issue is by introducing public health and vaccine knowledge programs and vaccines in areas with a lower education level. Chapter 3 also identifies that higher LST in a neighbourhood could influence COVID-19 risk. One way to address this is with policy tools to address the local socioeconomic conditions and modify urban LST through land use patterns, increased green density and urban design guidelines (Buyantuyev & Wu, 2010; Y. Zhang & Sun, 2019).

In Chapter 4, the highest seasonal risk in Spring and the lowest in summer may inform public health officials to prepare for peak seasons ahead. The spatiotemporal trends of the neighbourhoods are visualized in Section 4.4.3. The neighbourhoods that show increasing spatiotemporal risk over time could be the grounds to assess existing pandemic policies and can be further assessed by public health planners to observe the causes of the increasing patterns despite public health measures.

In chapter 5, the spread patterns and the sources of neighbourhood-scale outbreaks of COVID-19 are visualized in Section 5.4.2. During an epidemic, public health planners can apply the spatiotemporal spread to observe the movement of outbreaks across neighbourhoods in space and time to identify possible sources and directions of the outbreak. Focusing on the source neighbourhoods and the direction of the spread during the early stages of an outbreak could help public health planners to spread to further neighbourhoods with limited resources (Sharifi & Khavarian-Garmsir, 2020). This chapter also identifies the sources of the infections from the

dispersed neighbourhoods in Section 5.4.2.2. Such information can be useful for public health planners to understand the key sources of infections during a pandemic.

Neighbourhoods with high-risk clusters of COVID-19 in Toronto were visualized in Sections 3.4.2, 4.4.3 and 5.4.3. Given the limited resources available, planning efforts could focus on targeting high-risk neighbourhoods in Chapter 4 (Bayesian method), as observed in our dissertation.

6.3. Public Health Implications

This thesis has some major public health implications. Community-level interventions can play a significant role in reducing transmission and disease impact during an outbreak (Tiwari et al., 2021). It is important to protect individuals, especially vulnerable individuals, from the risk of severe disease outcomes or occupational exposure to the virus, which can reduce the probability of future variants to arise and reduce the pressure on health systems (van Seventer & Hochberg, 2017). According to the World Health Organization infectious disease outbreak control programs, public health efforts during an epidemic or pandemic should contain the following key strategies. First is the early detection and quick response to contain outbreaks (Murray & Cohen, 2017). The strategy is to focus on high-risk areas through early detection, and rapid response to contain the outbreaks is critical to reducing COVID-19 burdens (Franch-Pardo et al., 2020). Interventions like robust community engagement, strengthening the early warning systems, testing and vaccination capacities, health system and ready-for supplies while establishing the response teams at the core risk areas may drastically reduce the risk (Gilmore et al., 2020).

Several non-pharmaceutical mitigation and control strategies can be put in place at the community level to control the spread of the virus (Alvi et al., 2020; Odusanya et al., 2020). Non-pharmaceutical interventions are targeted to minimize the person-to-person spread of COVID-19 and can include bans on public gatherings, isolation, mandatory stay-at-home orders, mandating closures of schools and non-essential businesses, face mask ordinances, and quarantine at home or designated quarantine facilities where individuals are not allowed to leave (Hartley & Perencevich, 2020; Imai et al., 2020). These interventions can aim to prevent the viral spread and alleviate the effect in the short term. Particularly in the earliest phases of an outbreak, with the absence of pharmaceutical agents that are safe and effective, the public health community may have to rely on various non-pharmaceutical interventions to reduce the disease burdens of COVID-19 (Hartley & Perencevich, 2020).

Past studies have tested the effectiveness of these non-pharmaceutical interventions by analyzing historical observations (Fong et al., 2020; Hatchett et al., 2007; Markel et al., 2007). A common finding of these studies is that implementing non-pharmaceutical interventions of a newly contagious pathogen can significantly reduce transmission. Pan et al. (Pan et al., 2020) examined the epidemiologic outcomes following the implementation of non-pharmaceutical interventions during the COVID-19 outbreak in Wuhan, China. The results suggested that the series of multifaceted non-pharmaceutical interventions such as city lockdown with traffic restrictions, home quarantine, intensified social distancing measures, centralized quarantine and treatment, door to door and individual-to-individual community screening for symptoms in all residents and found that it was associated with improved control of COVID-19 outbreaks in Wuhan. The study also found that increased testing during the early phases of the outbreaks and

shelter-at-home- policies can work more effectively to replace citywide mass quarantine. It may be necessary to apply non-pharmaceutical interventions such as mass quarantine in the geographic hotspots identified in this thesis to limit the spread to other neighbourhoods in potential future waves of COVID-19. Such decisions can depend on factors such as the availability of testing or vaccines to measure population immunity accurately.

Second, a targeted multisectoral approach to targeting the COVID-19 hotspots identified in this study should be applied to improve prevention strategies. The spatial analytical methods employed in this study were able to identify hotspots which can be prioritized for interventions to reduce the COVID-19 infections identified in multiple hotspots in Toronto. The accessibility of hotspots supports the real-time formative evaluation of public health intervention during an ongoing pandemic (Hartley & Perencevich, 2020). Monitoring high-risk areas continuously may effectively allow for quality improvement methods to be used to evaluate public health policies (Butt et al., 2019). This thesis also identified some low-risk areas. From a public health standpoint, low-risk areas are also important as they could become high-risk areas if appropriate interventions to address the cause of the outbreak and higher transmission are not initiated in hotspots and in the neighbourhoods with consistently higher risk over the study period (Butt et al., 2019). Identifying high-risk neighbourhoods can enable health authorities to align their resources to reduce the overall COVID-19 disease burden and evaluate the impact of their prevention programmes. Over time, follow-up and surveillance of these high-risk areas are required to assess the impact of prevention programmes and activities (Maroko et al., 2020).

Additionally, this thesis identified some socioeconomic issues, such as neighbourhoods with a lower level of education having a higher risk of COVID-19, that may also contribute to

COVID-19 prevention. Core areas can be examined over time through awareness and can guide the COVID-19 public health programs. Contact tracing is another important element of the strategy to fight COVID-19 outbreaks that could be implemented. Contact tracing involves reaching out to close contacts of people with confirmed COVID-19 infections in the community and isolating those who have been in contact with people who have tested positive (Althoff et al., 2020). When public health units reach out to close contacts of confirmed cases to warn them of potential exposure and, in turn, take appropriate action to combat virus spread. As part of outbreak management, cases reaching a high level, public health must make a strategic shift to case and contact management and focus on the highest-risk areas (Steinbrook, 2020).

Pharmaceutical interventions, such as community-level vaccination in high-risk neighbourhoods, could also be an effective way to stop the infection (Milman et al., 2021). Prior studies have found that this strategy substantially speeds up and enhances the probability of containment (Chowell et al., 2019; Potluri et al., 2020). However, the success of this strategy depends on the level of community accessibility and cooperation (Milman et al., 2021). Interventions in the core risk areas, such as robust community engagement, establishing early warning surveillance, laboratory capacities, health systems, and establishing rapid response teams, can drastically reduce the number of deaths (Tabish, 2020).

Controlling any epidemics at the community level can require a well-implemented and relevant education of the members of the neighbourhoods with effective surveillance strategies integrated with the community. Furthermore, spatiotemporal spread, whether they are spreading and growing in magnitude, can help determine whether current mitigation and isolation techniques are effective in slowing the spread of COVID-19, suggesting that they could be used

more widely in public health (Coccia, 2020). For instance, in Italy, Giuliani et al. modelled the spatiotemporal spread of COVID-19 spread (Giuliani et al., 2020). The findings of the study showed strong evidence that strict control measures implemented in high-risk areas and monitoring the spread had effectively limited the spread to nearby areas. The spatiotemporal spread model in this study can be used to observe and monitor the magnitude of the spread in the neighbourhoods in Toronto to prevent future epidemics similar to COVID-19.

Finally, coordinating human, technical and financial resources in the hotspots could assist in selecting regions for targeted interventions to focus service delivery and resource allocation (Sarkar et al., 2020). In a pandemic such as COVID-19, the need for medical resources, including equipment and stuff, and space can quickly exceed available resources when allocating resources to combat this disease and to ensure that health programs and other interventions focus on the areas of greatest need (Hempel et al., 2020). Hotspots identified in this study can provide the basis for targeted disease control by optimizing resource allocation. An effective mechanism would be technical support, resource mobilization, partnership with local, provincial or global levels, supporting through human, technical, and financial resources (Djalante et al., 2020). Prioritizing the allocation of resources by focusing on areas with the highest disease risk as they have the greatest need for high priority in the allocation of resources that may include mobilizing resources, skilled human power or improved access to health facilities (Daniels, 2016).

6.4. Limitations

A geographic study design can often present some statistical challenges, and ignoring these challenges may result in potential bias in the estimated disease risk and the inference on the

ecological relationships (D. Lee & Sarran, 2015). For instance, in a spatial regression model, covariates as the confounding factors are often introduced to partially explain the variations in disease rates and to explain the relationship of the variable of interest with the outcome after adjusting for the covariates (Paciorek, 2010). However, getting all the confounding factors in the analysis is challenging. Thus, we often perform the analysis with these limitations. Also, some unidentified variables may also be missing, missing data patterns, or the data have been measured with errors that may be non-ignorable and could potentially affect our analysis (Dong & Peng, 2013). For example, the spatial patterns of travel behaviour during the pandemic may have varied across neighbourhoods which may have an impact on COVID-19 incidences. However, it is challenging to differentiate variations in mobility or travel patterns across the neighbourhoods overtime during the lockdown events. A future study including reliable mobility data at a finer spatiotemporal resolution could be explored.

The variations in healthcare provisions and unequal resource distributions may have impacted the regional variations in disease risk over the course of the pandemic in Toronto (Bambra et al., 2020; Clare et al., 2021) . There could also be a bias in the self-reporting information on income, education, housing conditions and other information (Ryu et al., 2018). A common confounding factor, such as age distribution in a small area, can sometimes have an impact on analysis (McCandless & Gustafson, 2017). However, this may not have been the case in this thesis as the coefficient of variation within an age group across all neighbourhoods was low (<10%), suggesting that the rates obtained under the age group distribution across the neighbourhoods had negligible impact on the analysis. It is also important to note that it may be possible that some interaction effects exist but were not detected due to possible insufficiency or

incorrectness of data or missing important confounding variables, which may act over different spatiotemporal scales (L. Liu et al., 2021).

Another limitation is the COVID-19 testing bias. This research includes only those individuals who did an official test which was recorded in the public health system. Unreported or asymptomatic cases are common problems in infectious disease research, as we may never know the true burden of the disease (Fernández-Fontelo et al., 2020). We assumed that the unreported and asymptomatic cases are randomly distributed across the neighbourhoods, thus, it did not affect much in our analysis. Furthermore, in the absence of restrictive measures such as lockdowns, travel restrictions, school openings and closing or mall closings, the entire population in our study area would have been exposed to the virus to more active cases and deaths (Shadmi et al., 2020). These various federal and local restrictive measures, such as the timing of lockdowns and school and restrictions of gatherings, may have varied over time in Toronto but may not have varied across the neighbourhoods in Toronto as the lockdowns were mandated for all of Toronto at the same time (Dainton & Hay, 2021).

In the statistical modelling of infectious diseases, one challenge is how to control for the spatial autocorrelation present in the data after accounting for the known covariates caused by the unmeasured confounders (Duncan et al., 2017; D. Lee & Sarran, 2015). Second is how to adjust the model to account for the spatial misalignment between disease and the covariates, which can cause the within area-variations in the covariate data. It is important to address these missing data or measurement errors and the confounders (Duncan et al., 2017; D. Lee & Sarran, 2015). For instance, if the spatially structured missing covariates are the confounders in the relationship, it may be difficult to assess the true impact, which is known as the problem of

spatial confounding (Paciorek, 2010). The generalized linear mixed models and Bayesian models that have been used in the study allowed for adjustment of the spatial confounding and could overcome the influences of these missing covariates (Jerrett et al., 2010). Spatially structured random and spatially unstructured random effects can help to cope with the effects of missing covariates in model specification (Haining & Li, 2020; Griffith, 2010; Law et al., 2006). It is suggested that spatial random effects in spatial GLMM models can explain the missing covariates and the unobserved heterogeneity in ecological analytical models (Cressie et al., 2009). A study performed a sensitivity analysis for ecological models and showed that the area-level random effects are a natural way to model the unmeasured confounders (Wakefield, 2003).

To account for the unobserved confounding factors and for the general missingness or errors in data, this thesis adopted a spatial GLMM model in the Bayesian framework that included structured and unstructured spatial and temporal random effects and a space-time interaction effect (Haining & Li, 2020; Lowe et al., 2011). Including these random effects introduced extra variables or a latent effect into the model and captured the impact of unknown and unobserved confounding factors (Lowe et al., 2013). Specifically, the Type II space–time interaction effect proposed by Knorr-Held (2000) used in the final Bayesian model in the thesis implied that the missing covariates have smoothly varying structures through time and have no structure over space since they are highly localized in their effect on the outcome (Knorr-Held, 2000; Haining & Li, 2020; Knorr-Held & Besag, 1998).

Another limitation in this study is that the data are aggregated at a neighbourhood level, thus, the results presented in this dissertation must be taken in the context of the neighbourhood scale and cannot be generalized at the individual level or at spatial units that are smaller or larger

(Sedgwick, 2015). Upon data availability, future research may focus on using point data to help addressing this limitation. Finally, the data on the other factors, such as pollution or health-seeking behaviour at the neighbourhood level, were unavailable, which could explain some of the complex relationships between urban environment and disease spread. One way to address this limitation is to collect these data at the neighbourhood level in future research. Chapters 3, 4, and 5 further addressed the specific limitations of each manuscript.

6.5. Future research

6.5.1. Continuing spatiotemporal analyses of the urban neighbourhoods

Future research can further explore the relationships between spatiotemporal COVID-19 patterns and urban environments, building contributions of this dissertation. Other models can also be explored to identify the spatiotemporal risk trends of COVID-19 and identify the risk factors. For example, future research can explore the spatially varying coefficient model (Finley & Banerjee, 2020), the Bayesian equivalent to the GWR model, that considers the non-stationarity of the coefficients, to assess the differences in findings of the results from our MGWR model and test for model accuracies. Additionally, new studies could develop and implement new statistical models to understand infectious disease epidemiology and spread patterns in an urban setting.

Furthermore, different countries or regions have different pandemic environments, demographic and government measures. The quantitative output of this thesis is not applicable equally to all countries or regions of the world. Therefore, future research is needed in other urban settings to understand the spatiotemporal trends and risks and relationships with the risk factors in relation to pandemic situations. Future research also needs more longitudinal studies, possibly expanding beyond several years after the COVID-19 pandemic.

6.5.2. Infectious disease modelling

Even after the end of this pandemic, it is necessary to continue to study COVID-19 studies to help understand the spread of other emerging or remerging infections that may come our way (Karmarkar et al., 2021; Nii-Trebi, 2017). Future research may apply the methodological framework of the applied and novel tools from this dissertation and apply in other infectious diseases. The recent COVID-19 pandemic exhibited that humans are not infallible, and communities need to be prepared for future infectious disease outbreaks (Khanna et al., 2020). A delay in detection and response has led to an overburdening health system in many regions around the world (Binny et al., 2021; Khanna et al., 2020). In future, with new emerging or remerging infectious diseases, these methods utilized in this dissertation could be incorporated into public health research at an early stage to control rapid spread.

A more urgent need for advancing the Bayesian Spatiotemporal models maybe to improve current methods of Bayesian estimation and inference (MacNab, 2022). INLA (Integrated Nested Laplace Approximation) may offer a solution to computational challenges for a large number of model parameters and provide alternative computational options for Monte Carlo sampling methods that can be further explored (MacNab, 2022).

6.5.3. Incorporating other explanatory variables

One potential avenue for future research could involve additional explanatory variables to model the risk of COVID-19 in urban settings. It would be insightful to collect and include socio-behavioural variables for each neighbourhood by the local government to explain some of the neighbourhood variations in COVID-19 transmission and trends (Bandhu Kalanidhi et al., 2021; Zajacova et al., 2020). The analysis of urban population mobility trends has the potential implications for future research to control COVID-19 outbreaks (Hu et al., 2021; Kraemer et al.,

2020; Rincón Ruiz et al., 2013). With the availability of population mobility data at a local level, future research could be studied that can provide more detailed information regarding COVID-19 transmissions and origin and further identify the most relevant mobilities associated with new infections and outbreaks. Furthermore, neighbourhood-level urban air pollution data can be collected by the local government and incorporated into predicting the risk of COVID-19 (Contini & Costabile, 2020; Travaglio et al., 2021).

Additionally, joint modelling of COVID-19 cases and deaths simultaneously while adjusting for covariates using Bayesian shared component models (Dreassi, 2013) could be useful for understanding the relationship between disease transmission and death in relation to the associated risk factors that influence the outcomes. Another important factor worth investigating is urban green space accessibility and use during pandemics across different socioeconomic groups (Kabisch & Haase, 2014; Kronenberg et al., 2020). Future studies incorporating individual or community-level data having temporal consistency with COVID-19 transmission time could produce better directives analyzing risk factors at a lower geographic unit and finding solutions for different socio-economic groups.

6.5.4. Public Health policy studies

The results of this dissertation could be used to assess the existing public health policies in future research. Future research can collect data related to vaccination rates, vaccine inequities, vaccine misinformation or vaccine hesitancy and incorporate them into the increasing or decreasing trends of COVID-19 to assess the efficacy of the vaccine programs (Mathieu et al., 2021; Suthar et al., 2022; Watson et al., 2022). The transmission mechanism of the pandemic may significantly differ from vaccinated to unvaccinated areas (Iftekhar et al., 2021). Future research can compare the transmission mechanism of the disease transmission in vaccinated and

unvaccinated areas. More studies are needed to evaluate the existing public health policy evaluations during a pandemic. Future research should focus on developing quantitative methods that study the effectiveness of different public health interventions, programs and policies to combat disease transmissions and deaths during the pandemic.

References

Aamodt, G., Samuelsen, S. O., & Skrondal, A. (2006). A simulation study of three methods for detecting disease clusters. *International Journal of Health Geographics*. https://doi.org/10.1186/1476-072X-5-15

Abedi, V., Olulana, O., Avula, V., Chaudhary, D., Khan, A., Shahjouei, S., Li, J., & Zand, R. (2020). Racial, Economic, and Health Inequality and COVID-19 Infection in the United States. *Journal of Racial and Ethnic Health Disparities*. https://doi.org/10.1007/s40615-020-00833-4

Acharya, B. K., Khanal, L., Mahyoub, A. S. M., Ruan, Z., Yang, Y., Adhikari, S. K., Pandit, S., Neupane, B. K., Paudel, B. K., Lin, H., Capobianchi, M. R., Agrifoglio, M., Colombo, G. I., Poggio, P., Pesce, M., Bruno, K. A., Paneth, N. S., Casadevall, A., Joyner, M. J., … Hensley, S. E. (2020). Execution of intervention matters more than strategy A lesson from the spatiotemporal assessment of COVID-19 clusters in Nepal. *MedRxiv*.

Ainsworth, L. M., & Dean, C. B. (2006). Approximate inference for disease mapping. *Computational Statistics & Data Analysis*, *50*(10), 2552–2570.

Alene, M., Yismaw, L., Assemie, M. A., Ketema, D. B., Mengist, B., Kassie, B., & Birhan, T. Y. (2021). Magnitude of asymptomatic COVID-19 cases throughout the course of infection: A systematic review and meta-analysis. *PLOS ONE*, *16*(3), e0249090. https://doi.org/10.1371/journal.pone.0249090

Ali, M., Emch, M., Yunus, M., & Clemens, J. (2009). Modeling spatial heterogeneity of disease risk and evaluation of the impact of vaccination. *Vaccine*. https://doi.org/10.1016/j.vaccine.2009.03.085

Alirol, E., Getaz, L., Stoll, B., Chappuis, F., & Loutan, L. (2011). Urbanisation and infectious diseases in a globalised world. *The Lancet Infectious Diseases*, *11*(2), 131–141. https://doi.org/10.1016/S1473-3099(10)70223-1

Alkhamis, M. A., Al Youha, S., Khajah, M. M., Ben Haider, N., Alhardan, S., Nabeel, A., Al Mazeedi, S., & Al-Sabah, S. K. (2020). Spatiotemporal dynamics of the COVID-19 pandemic in the State of Kuwait. *International Journal of Infectious Diseases*. https://doi.org/10.1016/j.ijid.2020.06.078

Althoff, K. N., Coburn, S. B., & Nash, D. (2020). Contact Tracing: Essential to the Public Health Response and Our Understanding of the Epidemiology of Coronavirus Disease 2019. *Clinical Infectious Diseases*, *71*(8), 1960–1961. https://doi.org/10.1093/cid/ciaa757

Alvi, M. M., Sivasankaran, S., & Singh, M. (2020). Pharmacological and non-pharmacological efforts at prevention, mitigation, and treatment for COVID-19. *Journal of Drug Targeting*, *28*(7–8), 742–754. https://doi.org/10.1080/1061186X.2020.1793990

Andersen, L. M., Harden, S. R., Sugg, M. M., Runkle, J. D., & Lundquist, T. E. (2021a). Analyzing the spatial determinants of local Covid-19 transmission in the United States. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.142396

Andersen, L. M., Harden, S. R., Sugg, M. M., Runkle, J. D., & Lundquist, T. E. (2021b). Analyzing the spatial determinants of local Covid-19 transmission in the United States. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.142396

Anderson, C. I., Gough, W. A., & Mohsin, T. (2018). Characterization of the urban heat island at Toronto: Revisiting the choice of rural sites using a measure of day-to-day variation. *Urban Climate*, *25*, 187–195. https://doi.org/10.1016/j.uclim.2018.07.002

Andrade, L. A., Gomes, D. S., Lima, S. V. M. A., Duque, A. M., Melo, M. S., Góes, M. A. O., Ribeiro, C. J. N., Peixoto, M. V. S., Souza, C. D. F., & Santos, A. D. (2020). COVID-19 Mortality in an area of northeast brazil: Epidemiological characteristics and prospective spatiotemporal modeling. *Epidemiology and Infection*. https://doi.org/10.1017/S0950268820002915

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

Anselin, L. (2018). *GeoDa Workbook*. Geoda WorkBook: Global Spatial Autocorrelation (1) - Moran Scatter Plot and Correlogram.

Anselin, L., Syabri, I., & Kho, Y. (2006). GeoDa: An introduction to spatial data analysis. *Geographical Analysis*. https://doi.org/10.1111/j.0016-7363.2005.00671.x

Anselin, L., Syabri, I., & Kho, Y. (2010). GeoDa: An Introduction to Spatial Data Analysis. In M. M. Fischer & A. Getis (Eds.), *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications* (pp. 73–89). Springer. https://doi.org/10.1007/978-3-642-03647-7_5

Arashi, M., Bekker, A., Salehi, M., Millard, S., Erasmus, B., Cronje, T., & Golpaygani, M. (2020). Spatial analysis and prediction of COVID-19 spread in South Africa after lockdown. In *ArXiv*.

Arcaya, M., Tucker-Seeley, R., Kim, R., Schnake-Mahl, A., So, M., & Subramanian, S. (2016). Research on Neighborhood Effects on Health in the United States: A Systematic Review of Study Characteristics. *Social Science & Medicine (1982)*, *168*, 16–29. https://doi.org/10.1016/j.socscimed.2016.08.047

Arif, M., & Sengupta, S. (2021). Nexus between population density and novel coronavirus (COVID-19) pandemic in the south Indian states: A geo-statistical approach. *Environment, Development and Sustainability*, *23*(7), 10246–10274. https://doi.org/10.1007/s10668-020-01055-8

Armitage, C. (2021). The high burden of infectious disease. *Nature*, *598*(7882), S9–S9. https://doi.org/10.1038/d41586-021-02909-5

Auchincloss, A. H., Gebreab, S. Y., Mair, C., & Diez Roux, A. V. (2012). A Review of Spatial Methods in Epidemiology 2000–2010. *Annual Review of Public Health*. https://doi.org/10.1146/annurev-publhealth-031811-124655

Azmach, N. N., Tesfahannes, T. G., Abdulsemed, S. A., & Hamza, T. A. (2020). *Prospective Time Periodic Geographical Covid-19 Surveillance in Ethiopia Using a Space-time Scan Statistics: Detecting and Evaluating Emerging Clusters*. Research Square. https://doi.org/10.21203/rs.3.rs-76052/v1

Badurdeen, S., Valladares, D. B., Farrar, J., Gozzer, E., Kroeger, A., Kuswara, N., Ranzinger, S. R., Tinh, H. T., Leite, P., Mahendradhata, Y., Skewes, R., & Verrall, A. (2013). Sharing experiences: Towards an evidence based model of dengue surveillance and outbreak response in Latin America and Asia. *BMC Public Health*, *13*(1), 607. https://doi.org/10.1186/1471-2458-13-607

Balcan, D., Gonçalves, B., Hu, H., Ramasco, J. J., Colizza, V., & Vespignani, A. (2010). Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. *Journal of Computational Science*, *1*(3), 132–145. https://doi.org/10.1016/j.jocs.2010.07.002

Ballesteros, P., Salazar, E., Sánchez, D., & Bolanos, C. (2020). Spatial and spatiotemporal clustering of the COVID-19 pandemic in Ecuador. *Revista de La Facultad de Medicina*. https://doi.org/10.15446/revfacmed.v69n1.86476

Bambra, C., Riordan, R., Ford, J., & Matthews, F. (2020). The COVID-19 pandemic and health inequalities. *J Epidemiol Community Health*, *74*(11), 964–968. https://doi.org/10.1136/jech-2020-214401

Bandhu Kalanidhi, K., Ranjan, P., Sarkar, S., Kaur, T., Dutt Upadhyay, A., Singh, A., Sahu, A., Khan, M., Vijay Prasad, B., Baitha, U., & Kumar, A. (2021). Development and validation of a questionnaire to assess socio-behavioural impact of COVID-19 on the general population. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *15*(2), 601–603. https://doi.org/10.1016/j.dsx.2021.02.019

Banos, A., & Lacasa, J. (2007). Spatio-temporal exploration of SARS epidemic. *Cybergeo: European Journal of Geography*. https://doi.org/10.4000/cybergeo.12803

Bayarri, M. J., Berger, J. O., Forte, A., & García-Donato, G. (2012). Criteria for Bayesian model choice with application to variable selection. *The Annals of Statistics*, *40*(3). https://doi.org/10.1214/12-AOS1013

Beale, L., Abellan, J. J., Hodgson, S., & Jarup, L. (2008). Methodologic issues and approaches to spatial epidemiology. *Environmental Health Perspectives*. https://doi.org/10.1289/ehp.10816

Beenstock, M., & Felsenstein, D. (2021). Freedom of Information and Personal Confidentiality in Spatial COVID-19 Data. *Journal of Official Statistics*, *37*(4), 791–809. https://doi.org/10.2478/jos-2021-0035

Béland, L.-P., Brodeur, A., Mikola, D., & Wright, T. (2020). *The Short-Term Economic Consequences of Covid-19: Occupation Tasks and Mental Health in Canada* (SSRN Scholarly Paper No. 3602430). Social Science Research Network. https://doi.org/10.2139/ssrn.3602430

Benita, F., & Gasca-Sanchez, F. (2021). The main factors influencing COVID-19 spread and deaths in Mexico: A comparison between phases I and II. *Applied Geography (Sevenoaks, England), 134*, 102523. https://doi.org/10.1016/j.apgeog.2021.102523

Benita, F., Gasca-Sanchez, F., Rosenbaum-Feldbruegge, M., Dudel, C., Loichinger, E., Sander, N., Backhaus, A., Fava, E. D., Esins, J., Fischer, M., Grabenhenrich, L., Grigoriev, P., Grow, A., Hilton, J., Koller, B., Myrskyla, M., Scalone, F., Wolkewitz, M., Zagheni, E., … Raoult, D. (2020). On the main factors influencing COVID-19 spread and deaths in Mexico A comparison between Phase I and II. *MedRxiv*.

Berke, O. (2004). Exploratory disease mapping: Kriging the spatial risk function from regional count data. *International Journal of Health Geographics*. https://doi.org/10.1186/1476-072X-3-18

Bermudi, P. M. M., Lorenz, C., Aguiar, B. S. de, Failla, M. A., Barrozo, L. V., & Chiaravalloti-Neto, F. (2021). Spatiotemporal ecological study of COVID-19 mortality in the city of São Paulo, Brazil: Shifting of the high mortality risk from areas with the best to those with the worst socio-economic conditions. *Travel Medicine and Infectious Disease, 39*, 101945. https://doi.org/10.1016/j.tmaid.2020.101945

Bernardinelli, L., & Montomoli, C. (1992). Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. *Statistics in Medicine*, *11*(8), 983–1007. https://doi.org/10.1002/sim.4780110802

Besag, J., York, J., & Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*. https://doi.org/10.1007/BF00116466

Best, N., Richardson, S., & Thomson, A. (2005). A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research*, *14*(1), 35–59. https://doi.org/10.1191/0962280205sm388oa

Bhadra, A., Mukherjee, A., & Sarkar, K. (2021). Impact of population density on Covid-19 infected and mortality rate in India. *Modeling Earth Systems and Environment*, *7*(1), 623–629. https://doi.org/10.1007/s40808-020-00984-7

Bie, S., Hu, X., Zhang, H., Wang, K., & Dou, Z. (2021). Influential factors and spatial-temporal distribution of tuberculosis in mainland China. *Scientific Reports*, *11*(1), 6274. https://doi.org/10.1038/s41598-021-85781-7

Bilal, U., Barber, S., & Diez-Roux, A. V. (2020). Spatial Inequities in COVID-19 outcomes in Three US Cities. *MedRxiv*.

Binder, S., Levitt, A. M., Sacks, J. J., & Hughes, J. M. (1999). Emerging Infectious Diseases: Public Health Issues for the 21st Century. *Science*, *284*(5418), 1311–1313. https://doi.org/10.1126/science.284.5418.1311

Binny, R. N., Baker, M. G., Hendy, S. C., James, A., Lustig, A., Plank, M. J., Ridings, K. M., & Steyn, N. (2021). Early intervention is the key to success in COVID-19 control. *Royal Society Open Science*, *8*(11), 210488. https://doi.org/10.1098/rsos.210488

Blangiardo, M., Boulieri, A., Diggle, P., Piel, F. B., Shaddick, G., & Elliott, P. (2020). Advances in spatiotemporal models for non-communicable disease surveillance. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dyz181

Blangiardo, M., Cameletti, M., Pirani, M., Corsetti, G., Battaglini, M., & Baio, G. (2020). Estimating weekly excess mortality at sub-national level in Italy during the COVID-19 pandemic. *PLOS ONE*. https://doi.org/10.1101/2020.06.08.20125211

Boland, M. R., Liu, J., Balocchi, C., Meeker, J., Bai, R., Mellis, I., Mowery, D. L., & Herman, D. (2021). Association of Neighborhood-Level Factors and COVID-19 Infection Patterns in Philadelphia Using Spatial Regression. *AMIA Summits on Translational Science Proceedings*, *2021*, 545–554.

Booth, A., Reed, A. B., Ponzo, S., Yassaee, A., Aral, M., Plans, D., Labrique, A., & Mohan, D. (2021). Population risk factors for severe disease and mortality in COVID-19: A global systematic review and meta-analysis. *PLOS ONE*, *16*(3), e0247461. https://doi.org/10.1371/journal.pone.0247461

Borjas, G. J. (2020). Demographic Determinants of Testing Incidence and COVID-19 Infections in New York City Neighborhoods. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3572329

Boulos, M. N. K. (2004). Descriptive review of geographic mapping of severe acute respiratory syndrome (SARS) on the Internet. *International Journal of Health Geographics*, *3*, 2. https://doi.org/10.1186/1476-072X-3-2

Brady, O. J., Smith, D. L., Scott, T. W., & Hay, S. I. (2015). Dengue disease outbreak definitions are implicitly variable. *Epidemics*, *11*, 92–102. https://doi.org/10.1016/j.epidem.2015.03.002

Briz-Redón, Á., Iftimi, A., Correcher, J. F., De Andrés, J., Lozano, M., & Romero-García, C. (2022). A comparison of multiple neighborhood matrix specifications for spatio-temporal model fitting: A case study on COVID-19 data. *Stochastic Environmental Research and Risk Assessment: Research Journal*, *36*(1), 271–282. https://doi.org/10.1007/s00477-021-02077-y

Briz-Redón, Á., & Serrano-Aroca, Á. (2020). A spatio-temporal analysis for exploring the effect of temperature on COVID-19 early evolution in Spain. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.138811

Brown, K. A., Soucy, J.-P. R., Buchan, S. A., Sturrock, S. L., Berry, I., Stall, N. M., Jüni, P., Ghasemi, A., Gibb, N., MacFadden, D. R., & Daneman, N. (2021). *The mobility gap: Estimating mobility levels required to control Canada's winter COVID-19 surge* (p. 2021.01.28.21250622). https://doi.org/10.1101/2021.01.28.21250622

Brunsdont, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression—Modelling spatial non-stationarity. *Journal of the Royal Statistical Society Series D: The Statistician*. https://doi.org/10.1111/1467-9884.00145

Bulut, C., & Kato, Y. (2020). Epidemiology of COVID-19. *Turkish Journal of Medical Sciences*, *50*(SI-1), 563–570. https://doi.org/10.3906/sag-2004-172

Burton, D. C., Flannery, B., Bennett, N. M., Farley, M. M., Gershman, K., Harrison, L. H., Lynfield, R., Petit, S., Reingold, A. L., Schaffner, W., Thomas, A., Plikaytis, B. D., Rose, C. E., Whitney, C. G., & Schuchat, A. (2010). Socioeconomic and Racial/Ethnic Disparities in the Incidence of Bacteremic Pneumonia Among US Adults. *American Journal of Public Health*, *100*(10), 1904–1911. https://doi.org/10.2105/AJPH.2009.181313

Butt, Z. A., Mak, S., Gesink, D., Gilbert, M., Wong, J., Yu, A., Wong, S., Alvarez, M., Chong, M., Buxton, J., Tyndall, M., Krajden, M., & Janjua, N. Z. (2019). Applying core theory and spatial analysis to identify hepatitis C virus infection 'core areas' in British Columbia, Canada. *Journal of Viral Hepatitis*, *26*(3), 373–383. https://doi.org/10.1111/jvh.13043

Buyantuyev, A., & Wu, J. (2010). Urban heat islands and landscape heterogeneity: Linking spatiotemporal variations in surface temperatures to land-cover and socioeconomic patterns. *Landscape Ecology*, *25*(1), 17–33. https://doi.org/10.1007/s10980-009-9402-4

Byun, H. G., Lee, N., & Hwang, S.-S. (2021). A Systematic Review of Spatial and Spatio-temporal Analyses in Public Health Research in Korea. *Journal of Preventive Medicine and Public Health = Yebang Uihakhoe Chi*, *54*(5), 301–308. https://doi.org/10.3961/jpmph.21.160

Cai, B., Lawson, A. B., Hossain, Md. M., Choi, J., Kirby, R. S., & Liu, J. (2013). Bayesian semiparametric model with spatially-temporally varying coefficients selection. *Statistics in Medicine, 32*(21), 3670–3685. https://doi.org/10.1002/sim.5789

Canada, P. H. A. of. (2021a, January 15). *Demographics: COVID-19 vaccination coverage in Canada - Canada.ca* [Datasets;statistics;education and awareness]. Aem. https://health-infobase.canada.ca/covid-19/vaccination-coverage/

Canada, P. H. A. of. (2021b, February 3). *National case definition: Coronavirus disease (COVID-19)* [Research]. https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/health-professionals/national-case-definition.html

Canada, P. H. A. of. (2022, June 27). *Summary of evidence supporting COVID-19 public health measures* [Transparency - other]. https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/guidance-documents/summary-evidence-supporting-covid-19-public-health-measures.html

Cao, Y., Hiyoshi, A., & Montgomery, S. (2020). COVID-19 case-fatality rate and demographic and socioeconomic influencers: Worldwide spatial regression analysis based on country-level data. *BMJ Open*. https://doi.org/10.1136/bmjopen-2020-043560

Caprarelli, G., & Fletcher, S. (2014). A brief review of spatial analysis concepts and tools used for mapping, containment and risk modelling of infectious diseases and other illnesses. *Parasitology, 141*(5), 581–601. https://doi.org/10.1017/S0031182013001972

Carroll, L. N., Au, A. P., Detwiler, L. T., Fu, T., Painter, I. S., & Abernethy, N. F. (2014). Visualization and analytics tools for infectious disease epidemiology: A systematic

review. *Journal of Biomedical Informatics, 51*, 287–298. https://doi.org/10.1016/j.jbi.2014.04.006

Castro, R. R., Santos, R. S. C., Sousa, G. J. B., Pinheiro, Y. T., Martins, R. R. I. M., Pereira, M. L. D., & Silva, R. A. R. (2021). Spatial dynamics of the COVID-19 pandemic in Brazil. *Epidemiology and Infection*, *149*, e60-undefined. https://doi.org/10.1017/S0950268821000479

CBC News. (2020). *Why it's so difficult to get tested for COVID-19 in Canada*. CBC. https://www.cbc.ca/news/health/covid-testing-shortages-1.5503926

CBC News. (2021). *Record high number of Ontarians in hospital with COVID-19 as Omicron wave continues*. https://www.cbc.ca/news/canada/toronto/covid-19-ontario-jan-7-2022-hospitalizaitons-pandemic-high-1.6307179

CDC. (2021, July 19). *CDC SARS Response Timeline | About | CDC*. https://www.cdc.gov/about/history/sars/timeline.htm

Chakraborty, J. (2021). Social inequities in the distribution of COVID-19: An intra-categorical analysis of people with disabilities in the U.S. *Disability and Health Journal*. https://doi.org/10.1016/j.dhjo.2020.101007

Chaudhry, R., Dranitsaris, G., Mubashir, T., Bartoszko, J., & Riazi, S. (2020). A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine*. https://doi.org/10.1016/j.eclinm.2020.100464

Chaudhuri, K., Chakrabarti, A., Lima, J. M., Chandan, J. S., & Bandyopadhyay, S. (2021). The interaction of ethnicity and deprivation on COVID-19 mortality risk: A retrospective

ecological study. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-91076-8

Chen, D. (2014). Modeling the Spread of Infectious Diseases: A Review. In *Analyzing and Modeling Spatial and Temporal Dynamics of Infectious Diseases* (pp. 19–42). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118630013.ch2

Chen, J., Roth, R. E., Naito, A. T., Lengerich, E. J., & MacEachren, A. M. (2008). Geovisual analytics to enhance spatial scan statistic interpretation: An analysis of U.S. cervical cancer mortality. *International Journal of Health Geographics*. https://doi.org/10.1186/1476-072X-7-57

Chen, W.-H., Myles, J., & Picot, G. (2012). Why Have Poorer Neighbourhoods Stagnated Economically while the Richer Have Flourished? Neighbourhood Income Inequality in Canadian Cities. *Urban Studies*, *49*(4), 877–896. https://doi.org/10.1177/0042098011408142

Chen, X., & Quan, R. (2021). A spatiotemporal analysis of urban resilience to the COVID-19 pandemic in the Yangtze River Delta. *Natural Hazards*, *106*(1), 829–854. https://doi.org/10.1007/s11069-020-04493-9

Chen, Y., & Jiao, J. (2020a). Relationship between Socio-Demographics and COVID-19: A Case Study in Three Texas Regions. *Ssrn*.

Chen, Y., & Jiao, J. (2020b). Relationship between Socio-Demographics and COVID-19: A Case Study in Three Texas Regions. *Ssrn*.

Chen, Y., Zhu, M., Zhou, Q., & Qiao, Y. (2021). Research on Spatiotemporal Differentiation and Influence Mechanism of Urban Resilience in China Based on MGWR Model.

International Journal of Environmental Research and Public Health, *18*(3), Article 3. https://doi.org/10.3390/ijerph18031056

Chien, L.-C., Chen, L.-W. A., & Lin, R.-T. (2021). Lagged meteorological impacts on COVID-19 incidence among high-risk counties in the United States—A spatiotemporal analysis. *Journal of Exposure Science & Environmental Epidemiology*, 1–8. https://doi.org/10.1038/s41370-021-00356-y

Cho, S., Lambert, D. M., Kim, S. G., & Jung, S. (2009). Extreme coefficients in geographically weighted regression and their effects on mapping. *GIScience and Remote Sensing*. https://doi.org/10.2747/1548-1603.46.3.273

Choi, K. H., Denice, P., Haan, M., & Zajacova, A. (2020). *Studying the social determinants of COVID-19 in a data vacuum*. SocArXiv. https://doi.org/10.31235/osf.io/yq8vu

Choi, K. H., Denice, P., Haan, M., & Zajacova, A. (2021). Studying the social determinants of COVID-19 in a data vacuum. *Canadian Review of Sociology/Revue Canadienne de Sociologie*, *58*(2), 146–164. https://doi.org/10.1111/cars.12336

Chow, T. E., Choi, Y., Yang, M., Mills, D., & Yue, R. (2020). Geographic pattern of human mobility and COVID-19 before and after Hubei lockdown. *Annals of GIS*. https://doi.org/10.1080/19475683.2020.1841828

Chowell, G., & Rothenberg, R. (2018a). Spatial infectious disease epidemiology: On the cusp. *BMC Medicine*. https://doi.org/10.1186/s12916-018-1184-6

Chowell, G., & Rothenberg, R. (2018b). Spatial infectious disease epidemiology: On the cusp. In *BMC Medicine*. https://doi.org/10.1186/s12916-018-1184-6

Chowell, G., Tariq, A., & Kiskowski, M. (2019). Vaccination strategies to control Ebola epidemics in the context of variable household inaccessibility levels. *PLOS Neglected Tropical Diseases, 13*(11), e0007814. https://doi.org/10.1371/journal.pntd.0007814

Chua, P. L. C., Ng, C. F. S., Tobias, A., Seposo, X. T., & Hashizume, M. (2022). Associations between ambient temperature and enteric infections by pathogen: A systematic review and meta-analysis. *The Lancet Planetary Health*, *6*(3), e202–e218. https://doi.org/10.1016/S2542-5196(22)00003-1

City of Toronto. (2017, November 14). *Neighbourhood Profiles* (Toronto, Ontario, Canada). City of Toronto; City of Toronto. https://www.toronto.ca/city-government/data-research-maps/neighbourhoods-communities/neighbourhood-profiles/

City of Toronto. (2020a). *Coronavirus disease 2019 (COVID-19): Epidemiology update*. https://www.toronto.ca/home/covid-19/covid-19-latest-city-of-toronto-news/covid-19-status-of-cases-in-toronto/

City of Toronto. (2020b). *Toronto Neighbourhood Profiles*. https://www.toronto.ca/city-government/data-research-maps/neighbourhoods-communities/neighbourhood-profiles/

City of Toronto. (2021a). *Toronto at a Glance*. https://www.toronto.ca/city-government/data-research-maps/toronto-at-a-glance/#:~:text=Located%20on%20a%20broad%20sloping,perimeter%20is%20approximately%20180%20km.

City of Toronto. (2021b, January 18). *COVID-19: Where to Get Vaccinated* (Toronto, Ontario, Canada). City of Toronto; City of Toronto. https://www.toronto.ca/home/covid-19/covid-19-vaccines/covid-19-how-to-get-vaccinated/

City of Toronto. (2022a). *COVID-19 Cases in Toronto*. Open Data Catelogue. https://open.toronto.ca/dataset/covid-19-cases-in-toronto/

City of Toronto. (2022b, January 4). *City of Toronto and its community partners continue to support at-risk communities in fight against latest wave of COVID-19* (Toronto, Ontario, Canada). City of Toronto; City of Toronto. https://www.toronto.ca/news/city-of-toronto-and-its-community-partners-continue-to-support-at-risk-communities-in-fight-against-latest-wave-of-covid-19/

Clare, B., Julia, L., & Smith, K. E. (2021). *The Unequal Pandemic: COVID-19 and Health Inequalities*. Policy Press.

Clark, E., Fredricks, K., Woc-Colburn, L., Bottazzi, M. E., & Weatherhead, J. (2020). Disproportionate impact of the COVID-19 pandemic on immigrant communities in the United States. *PLOS Neglected Tropical Diseases*, *14*(7), e0008484. https://doi.org/10.1371/journal.pntd.0008484

Clarke, K. C., McLafferty, S. L., & Tempalski, B. J. (1996). On Epidemiology and Geographic Information Systems: A Review and Discussion of Future Directions. *Emerging Infectious Diseases*. https://doi.org/10.3201/eid0202.960202

Clayton, D., & Bernardinelli, L. (1996). 205Bayesian methods for mapping disease risk. In P. Elliott, J. Cuzick, D. English, & R. Stern (Eds.), *Geographical and Environmental Epidemiology: Methods for Small Area Studies* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780192622358.003.0018

Coccia, M. (2020). Factors determining the diffusion of COVID-19 and suggested strategy to prevent future accelerated viral infectivity similar to COVID. *The Science of the Total Environment, 729*, 138474. https://doi.org/10.1016/j.scitotenv.2020.138474

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Coleman, M., Coleman, M., Mabuza, A. M., Kok, G., Coetzee, M., & Durrheim, D. N. (2009). Using the SaTScan method to detect local malaria clusters for guiding malaria control programmes. *Malaria Journal*. https://doi.org/10.1186/1475-2875-8-68

Coll, C., & Caselles, V. (1997). A split-window algorithm for land surface temperature from advanced very high resolution radiometer data: Validation and algorithm comparison. *Journal of Geophysical Research: Atmospheres*, *102*(D14), 16697–16713. https://doi.org/10.1029/97JD00929

Contini, D., & Costabile, F. (2020). Does Air Pollution Influence COVID-19 Outbreaks? *Atmosphere*, *11*(4), Article 4. https://doi.org/10.3390/atmos11040377

Cooper, D. L., Smith, G. E., Regan, M., Large, S., & Groenewegen, P. P. (2008). Tracking the spatial diffusion of influenza and norovirus using telehealth data: A spatiotemporal analysis of syndromic data. *BMC Medicine*, *6*(1), 16. https://doi.org/10.1186/1741-7015-6-16

Cordes, J., & Castro, M. C. (2020). Spatial analysis of COVID-19 clusters and contextual factors in New York City. *Spatial and Spatio-Temporal Epidemiology*. https://doi.org/10.1016/j.sste.2020.100355

Corpas-Burgos, F., García-Donato, G., & Martinez-Beneito, M. A. (2018). Some findings on zero-inflated and hurdle poisson models for disease mapping. *Statistics in Medicine*, *37*(23), 3325–3337. https://doi.org/10.1002/sim.7819

COVID-19 Tracker Canada. (2021). https://covid19tracker.ca/

Cox, F. E. G., & Bia, F. J. (2017). Historical overview of global infectious diseases and geopolitics. In *Infectious Diseases* (pp. 1–11). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119085751.ch1

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., & Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, *19*(3), 553–570. https://doi.org/10.1890/07-0744.1

CTV. (2020). *COVID-19 disproportionately impacted immigrants and refugees in Ontario, new report finds | CTV News*. https://toronto.ctvnews.ca/covid-19-disproportionately-impacted-immigrants-and-refugees-in-ontario-new-report-finds-1.5097363

CTV News. (2021). *COVID-19 Ontario: Record 18,445 new infections reported amid new testing guidelines*. https://toronto.ctvnews.ca/ontario-breaks-single-day-record-once-again-with-more-than-18-000-new-covid-19-cases-1.5725202

Cuadros, D. F., Xiao, Y., Mukandavire, Z., Correa-Agudelo, E., Hernández, A., Kim, H., & MacKinnon, N. J. (2020). Spatiotemporal transmission dynamics of the COVID-19 pandemic and its impact on critical healthcare capacity. *Health and Place*. https://doi.org/10.1016/j.healthplace.2020.102404

Cucinotta, D., & Vanelli, M. (2020). WHO Declares COVID-19 a Pandemic. *Acta Bio-Medica: Atenei Parmensis*, *91*(1), 157–160. https://doi.org/10.23750/abm.v91i1.9397

Cudahy, P. G. T., Andrews, J. R., Bilinski, A., Dowdy, D. W., Mathema, B., Menzies, N. A., Salomon, J. A., Shrestha, S., & Cohen, T. (2019). Spatially targeted screening to reduce tuberculosis transmission in high-incidence settings. *The Lancet Infectious Diseases*, *19*(3), e89–e95. https://doi.org/10.1016/S1473-3099(18)30443-2

Cullen, J. R., Chitprarop, U., Doberstyn, E. B., & Sombatwattanangkul, K. (1984). An epidemiological early warning system for malaria control in northern Thailand. *Bulletin of the World Health Organization*, *62*(1), 107–114.

Curtis, A., Ajayakumar, J., Curtis, J., Mihalik, S., Purohit, M., Scott, Z., Muisyo, J., Labadorf, J., Vijitakula, S., Yax, J., & Goldberg, D. W. (2020). Geographic monitoring for early disease detection (GeoMEDD). *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-78704-5

Curtis, A. J., Ajayakumar, J., Curtis, J., & Brown, S. (2022). Spatial Syndromic Surveillance and COVID-19 in the U.S.: Local Cluster Mapping for Pandemic Preparedness. *International Journal of Environmental Research and Public Health*, *19*(15), Article 15. https://doi.org/10.3390/ijerph19158931

Cutler, D. M., & Summers, L. H. (2020). The COVID-19 Pandemic and the $16 Trillion Virus. *JAMA*, *324*(15), 1495–1496. https://doi.org/10.1001/jama.2020.19759

Czabanowska, K., & Kuhlmann, E. (2021). Public health competences through the lens of the COVID-19 pandemic: What matters for health workforce preparedness for global health emergencies. *The International Journal of Health Planning and Management*, *36*(S1), 14–19. https://doi.org/10.1002/hpm.3131

da Costa, A. C. C., Codeço, C. T., Krainski, E. T., da Costa Gomes, M. F., & Nobre, A. A. (2018). Spatiotemporal diffusion of influenza A (H1N1): Starting point and risk factors. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0202832

da Silva, A. R., & Fotheringham, A. S. (2016). The Multiple Testing Issue in Geographically Weighted Regression. *Geographical Analysis*, *48*(3), 233–247. https://doi.org/10.1111/gean.12084

Da Silveira Moreira, R. (2020). COVID-19: Intensive care units, mechanical ventilators, and latent mortality profiles associated with case-fatality in Brazil. *Cadernos de Saude Publica*. https://doi.org/10.1590/0102-311X00080020

Dainton, C., & Hay, A. (2021). Quantifying the relationship between lockdowns, mobility, and effective reproduction number (Rt) during the COVID-19 pandemic in the Greater Toronto Area. *BMC Public Health*, *21*(1), 1658. https://doi.org/10.1186/s12889-021-11684-x

Daniels, N. (2016). Resource Allocation and Priority Setting. In *Public Health Ethics: Cases Spanning the Globe [Internet]*. Springer. https://doi.org/10.1007/978-3-319-23847-0_3

Das, A., Ghosh, S., Das, K., Basu, T., Dutta, I., & Das, M. (2020). Living environment matters: Unravelling the spatial clustering of COVID-19 hotspots in Kolkata megacity, India. *Sustainable Cities and Society*. https://doi.org/10.1016/j.scs.2020.102577

De Ridder, D., Sandoval, J., Vuilleumier, N., Azman, A. S., Stringhini, S., Kaiser, L., Joost, S., & Guessous, I. (2021). Socioeconomically Disadvantaged Neighborhoods Face Increased Persistence of SARS-CoV-2 Clusters. *Frontiers in Public Health*, *8*, 626090. https://doi.org/10.3389/fpubh.2020.626090

Debes, A. K., Shaffer, A. M., Ndikumana, T., Liesse, I., Ribaira, E., Djumo, C., Ali, M., & Sack, D. A. (2021). Cholera Hot-Spots and Contextual Factors in Burundi, Planning for Elimination. *Tropical Medicine and Infectious Disease*, *6*(2), Article 2. https://doi.org/10.3390/tropicalmed6020076

Deguen, S., & Kihal-Talantikite, W. (2021). Geographical Pattern of COVID-19-Related Outcomes over the Pandemic Period in France: A Nationwide Socio-Environmental Study. *International Journal of Environmental Research and Public Health*, *18*(4), 1824. https://doi.org/10.3390/ijerph18041824

del Rio-Chanona, R. M., Mealy, P., Pichler, A., Lafond, F., & Farmer, J. D. (2020). Supply and demand shocks in the COVID-19 pandemic: An industry and occupation perspective. *Oxford Review of Economic Policy*, *36*(Supplement_1), S94–S137. https://doi.org/10.1093/oxrep/graa033

Demena, B. A., Floridi, A., & Wagner, N. (2022). The Short-Term Impact of COVID-19 on LabourLabour Market Outcomes: Comparative Systematic Evidence. In E. Papyrakis (Ed.), *COVID-19 and International Development* (pp. 71–88). Springer International Publishing. https://doi.org/10.1007/978-3-030-82339-9_6

Demenech, L. M., Dumith, S. de C., Vieira, M. E. C. D., & Neiva-Silva, L. (2020). Income inequality and risk of infection and death by COVID-19 in Brazil. *Revista Brasileira de Epidemiologia = Brazilian Journal of Epidemiology*. https://doi.org/10.1590/1980-549720200095

Desjardins, M. R., Hohl, A., & Delmelle, E. M. (2020). Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. *Applied Geography*. https://doi.org/10.1016/j.apgeog.2020.102202

Desmet, K., & Wacziarg, R. (2020). Understanding Spatial Variation in COVID-19 across the United States. *NBER Working Paper Series*.

Detsky, A. S., & Bogoch, I. I. (2020). COVID-19 in Canada: Experience and Response. *JAMA - Journal of the American Medical Association*. https://doi.org/10.1001/jama.2020.14033

DiMaggio, C., Klein, M., Berry, C., & Frangos, S. (2020). Black/African American Communities are at highest risk of COVID-19: Spatial modeling of New York City ZIP Code–level testing results. *Annals of Epidemiology*. https://doi.org/10.1016/j.annepidem.2020.08.012

Djalante, R., Nurhidayah, L., Van Minh, H., Phuong, N. T. N., Mahendradhata, Y., Trias, A., Lassa, J., & Miller, M. A. (2020). COVID-19 and ASEAN responses: Comparative policy analysis. *Progress in Disaster Science*, *8*, 100129. https://doi.org/10.1016/j.pdisas.2020.100129

do Nascimento, A. C. L., Galvani, E., Gobo, J. P. A., & Wollmann, C. A. (2022). Comparison between Air Temperature and Land Surface Temperature for the City of São Paulo, Brazil. *Atmosphere*, *13*(3), Article 3. https://doi.org/10.3390/atmos13030491

Dom, N. C., Ahmad, A. H., Adawiyah, R., & Ismail, R. (2010). Spatial mapping of temporal risk characteristic of dengue cases in Subang Jaya. *2010 International Conference on Science and Social Research (CSSR 2010)*, 361–366. https://doi.org/10.1109/CSSR.2010.5773800

Dong, Y., & Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, *2*, 222. https://doi.org/10.1186/2193-1801-2-222

Dongmei Chen, Bernard Moulin, & Jianhong Wu. (2014). *Analyzing and Modeling Spatial and Temporal Dynamics of Infectious Diseases | Wiley Series in Probability and Statistics*. https://onlinelibrary.wiley.com/doi/book/10.1002/9781118630013

Donša, D., Grujić, V. J., Pipenbaher, N., & Ivajnšič, D. (2021). The Lyme Borreliosis Spatial Footprint in the 21st Century: A Key Study of Slovenia. *International Journal of Environmental Research and Public Health*, *18*(22), Article 22. https://doi.org/10.3390/ijerph182212061

Dreassi, E. (2013). Shared Components Models in Joint Disease Mapping: A Comparison. In A. Giusti, G. Ritter, & M. Vichi (Eds.), *Classification and Data Mining* (pp. 207–214). Springer. https://doi.org/10.1007/978-3-642-28894-4_25

Du, J., Song, K., & Yan, B. (2019). Impact of the Zhalong Wetland on Neighboring Land Surface Temperature Based on Remote Sensing and GIS. *Chinese Geographical Science*, *29*(5), 798–808. https://doi.org/10.1007/s11769-019-1050-2

Du, Z., Wang, Z., Liu, Y., Wang, H., Xue, F., & Liu, Y. (2014). Ecological niche modeling for predicting the potential risk areas of severe fever with thrombocytopenia syndrome. *International Journal of Infectious Diseases*, *26*, 1–8. https://doi.org/10.1016/j.ijid.2014.04.006

Dummer, T. J. B. (2008). Health geography: Supporting public health policy and planning. *CMAJ*. https://doi.org/10.1503/cmaj.071783

Duncan, E. W., White, N. M., & Mengersen, K. (2017). Spatial smoothing in Bayesian models: A comparison of weights matrix specifications and their impact on inference. *International Journal of Health Geographics*, *16*(1), 47. https://doi.org/10.1186/s12942-017-0120-x

Ejiagha, I. R., Ahmed, M. R., Hassan, Q. K., Dewan, A., Gupta, A., & Rangelova, E. (2020). Use of Remote Sensing in Comprehending the Influence of Urban Landscape's Composition and Configuration on Land Surface Temperature at Neighbourhood Scale. *Remote Sensing*, *12*(15), Article 15. https://doi.org/10.3390/rs12152508

Elliot, P., Wakefield, J. C., Best, N. G., & Briggs, D. J. (2000). Spatial epidemiology: Methods and applications. *Spatial Epidemiology: Methods and Applications.* https://www.cabdirect.org/cabdirect/abstract/20023007010

Elliott, P., & Wartenberg, D. (2004). Spatial epidemiology: Current approaches and future challenges. *Environmental Health Perspectives*. https://doi.org/10.1289/ehp.6735

Escobar, G. J., Adams, A. S., Liu, V. X., Soltesz, L., Chen, Y.-F. I., Parodi, S. M., Ray, G. T., Myers, L. C., Ramaprasad, C. M., Dlott, R., & Lee, C. (2021). Racial Disparities in COVID-19 Testing and Outcomes: Retrospective Cohort Study in an Integrated Health System. *Annals of Internal Medicine*, *174*(6), 786–793. https://doi.org/10.7326/M20-6979

Etxeberria, J., Goicoa, T., & Ugarte, M. D. (2018). Joint modelling of brain cancer incidence and mortality using Bayesian age- and gender-specific shared component models. *Stochastic Environmental Research and Risk Assessment*, *32*(10), 2951–2969. https://doi.org/10.1007/s00477-018-1567-4

F. Dormann, C., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., M. Schurr, F., & Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography*. https://doi.org/10.1111/j.2007.0906-7590.05171.x

Feinhandler, I., Cilento, B., Beauvais, B., Harrop, J., & Fulton, L. (2020). Predictors of Death Rate during the COVID-19 Pandemic. *Healthcare*. https://doi.org/10.3390/healthcare8030339

Feng, C. (2021). Spatial-temporal generalized additive model for modeling COVID-19 mortality risk in Toronto, Canada. *Spatial Statistics*, 100526. https://doi.org/10.1016/j.spasta.2021.100526

Feng, Y., Li, Q., Tong, X., Wang, R., Zhai, S., Gao, C., Lei, Z., Chen, S., Zhou, Y., Wang, J., Yan, X., Xie, H., Chen, P., Liu, S., Xv, X., Liu, S., Jin, Y., Wang, C., Hong, Z., … Guo, Y. (2020). Spatiotemporal spread pattern of the COVID-19 cases in China. *PLOS ONE*, *15*(12), e0244351. https://doi.org/10.1371/journal.pone.0244351

Fernández-Fontelo, A., Moriña, D., Cabaña, A., Arratia, A., & Puig, P. (2020). Estimating the real burden of disease under a pandemic situation: The SARS-CoV2 case. *PLOS ONE*, *15*(12), e0242956. https://doi.org/10.1371/journal.pone.0242956

Ferreira, M. C. (2020). Spatial association between the incidence rate of Covid-19 and poverty in the São Paulo municipality, Brazil. *Geospatial Health*. https://doi.org/10.4081/gh.2020.921

Ferreira, R. V., Martines, M. R., Toppa, R. H., Assunção, L. M., Desjardins, M. R., & Delmelle, E. M. (2020). Applying a prospective space-time scan statistic to examine the evolution of COVID-19 clusters in the state of São Paulo, Brazil. In *MedRxiv*. https://doi.org/10.1101/2020.06.04.20122770

Fielding-Miller, R., Sundaram, M., & Brouwer, K. (2020). Social determinants of COVID-19 mortality at the county level. *PLOS ONE*. https://doi.org/10.1101/2020.05.03.20089698

Finger, F., Bertuzzo, E., Luquero, F. J., Naibei, N., Touré, B., Allan, M., Porten, K., Lessler, J., Rinaldo, A., & Azman, A. S. (2018). The potential impact of case-area targeted interventions in response to cholera outbreaks: A modeling study. *PLOS Medicine*, *15*(2), e1002509. https://doi.org/10.1371/journal.pmed.1002509

Finley, A. O. (2011). Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence: *Spatially-varying coefficients models*. *Methods in Ecology and Evolution*, *2*(2), 143–154. https://doi.org/10.1111/j.2041-210X.2010.00060.x

Finley, A. O., & Banerjee, S. (2020). Bayesian spatially varying coefficient models in the spBayes R package. *Environmental Modelling & Software*, *125*, 104608. https://doi.org/10.1016/j.envsoft.2019.104608

Fletcher-Lartey, S. M., & Caprarelli, G. (2016). Application of GIS technology in public health: Successes and challenges. *Parasitology*, *143*(4), 401–415. https://doi.org/10.1017/S0031182015001869

Fong, M. W., Gao, H., Wong, J. Y., Xiao, J., Shiu, E. Y. C., Ryu, S., & Cowling, B. J. (2020). Nonpharmaceutical Measures for Pandemic Influenza in Nonhealthcare Settings-Social

Distancing Measures. *Emerging Infectious Diseases*, *26*(5), 976–984. https://doi.org/10.3201/eid2605.190995

Fotheringham, A. S., Yang, W., & Kang, W. (2017). Multiscale Geographically Weighted Regression (MGWR). *Annals of the American Association of Geographers*, *107*(6), 1247–1265. https://doi.org/10.1080/24694452.2017.1352480

Fotiadis, A., Polyzos, S., & Huan, T.-C. T. C. (2021). The good, the bad and the ugly on COVID-19 tourism recovery. *Annals of Tourism Research*, *87*, 103117. https://doi.org/10.1016/j.annals.2020.103117

Franch-Pardo, I., Desjardins, M. R., Barea-Navarro, I., & Cerdà, A. (2021). A review of GIS methodologies to analyze the dynamics of COVID-19 in the second half of 2020. *Transactions in GIS*, *25*(5), 2191–2239. https://doi.org/10.1111/tgis.12792

Franch-Pardo, I., Napoletano, B. M., Rosete-Verges, F., & Billa, L. (2020). Spatial analysis and GIS in the study of COVID-19. A review. *Science of The Total Environment*, *739*, 140033. https://doi.org/10.1016/j.scitotenv.2020.140033

Frieden, T. R. (2010). A Framework for Public Health Action: The Health Impact Pyramid. *American Journal of Public Health*, *100*(4), 590–595. https://doi.org/10.2105/AJPH.2009.185652

Fullilove, M. T. (2003). Neighborhoods and Infectious Diseases. In I. Kawachi & L. F. Berkman (Eds.), *Neighborhoods and Health* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195138382.003.0009

Galvani, A. P., Lei, X., & Jewell, N. P. (2003). Severe Acute Respiratory Syndrome: Temporal Stability and Geographic Variation in Death Rates and Doubling Times. *Emerging Infectious Diseases, 9*(8), 991–994. https://doi.org/10.3201/eid0908.030334

Ganegoda, N. C., Wijaya, K. P., Amadi, M., Erandi, K. K. W. H., & Aldila, D. (2021). Interrelationship between daily COVID-19 cases and average temperature as well as relative humidity in Germany. *Scientific Reports*, *11*(1), 11302. https://doi.org/10.1038/s41598-021-90873-5

Gao, X., Li, G., Wang, J., & Xu, T. (2021). Spatiotemporal evolution, pattern of diffusion, and influencing factors of the COVID-19 epidemic in Hainan Province, China. *Journal of Medical Virology*. https://doi.org/10.1002/jmv.27502

Garcia-Donato, G., & Forte, A. (2018). Bayesian Testing, Variable Selection and Model Averaging in Linear Models using R with BayesVarSel. *The R Journal*, *10*(1), 155. https://doi.org/10.32614/RJ-2018-021

Gayawan, E., Awe, O. O., Oseni, B. M., Uzochukwu, I. C., Adekunle, A., Samuel, G., Eisen, D. P., & Adegboye, O. A. (2020). The spatio-temporal epidemic dynamics of COVID-19 outbreak in Africa. *Epidemiology & Infection*, *148*. https://doi.org/10.1017/S0950268820001983

Gelfand, A. E., Kim, H.-J., Sirmans, C. F., & Banerjee, S. (2003). Spatial Modeling With Spatially Varying Coefficient Processes. *Journal of the American Statistical Association, 98*(462), 387–396. https://doi.org/10.1198/016214503000170

Gesler, W. (1986). The uses of spatial analysis in medical geography: A review. *Social Science and Medicine*. https://doi.org/10.1016/0277-9536(86)90253-4

Ghosh, P., & Cartone, A. (2020). A Spatio-temporal analysis of COVID-19 outbreak in Italy. *Regional Science Policy & Practice*, *12*(6), 1047–1062. https://doi.org/10.1111/rsp3.12376

Gianquintieri, L., Brovelli, M. A., Pagliosa, A., Dassi, G., Brambilla, P. M., Bonora, R., Sechi, G. M., & Caiani, E. G. (2020). Mapping spatiotemporal diffusion of COVID-19 in Lombardy (Italy) on the base of emergency medical services activities. *ISPRS International Journal of Geo-Information*. https://doi.org/10.3390/ijgi9110639

Gilmore, B., Ndejjo, R., Tchetchia, A., de Claro, V., Mago, E., Diallo, A. A., Lopes, C., & Bhattacharyya, S. (2020). Community engagement for COVID-19 prevention and control: A rapid evidence synthesis. *BMJ Global Health*, *5*(10), e003188. https://doi.org/10.1136/bmjgh-2020-003188

Giuliani, D., Dickson, M. M., Espa, G., & Santi, F. (2020). Modelling and predicting the spatio-temporal spread of Coronavirus disease 2019 (COVID-19) in Italy. *ArXiv*. https://doi.org/10.2139/ssrn.3559569

Gog, J. R., Ballesteros, S., Viboud, C., Simonsen, L., Bjornstad, O. N., Shaman, J., Chao, D. L., Khan, F., & Grenfell, B. T. (2014). Spatial Transmission of 2009 Pandemic Influenza in the US. *PLoS Computational Biology*, *10*(6). https://doi.org/10.1371/journal.pcbi.1003635

Goicoa, T., Ugarte, M. D., Etxeberria, J., & Militino, A. F. (2016). Age–space–time CAR models in Bayesian disease mapping. *Statistics in Medicine*, *35*(14), 2391–2405. https://doi.org/10.1002/sim.6873

Gomes, D. S., Andrade, L. A., Ribeiro, C. J. N., Peixoto, M. V. S., Lima, S. V. M. A., Duque, A. M., Cirilo, T. M., Goés, M. A. O., Lima, A. G. C. F., Santos, M. B., Araújo, K. C. G. M., & Santos, A. D. (2020). Risk clusters of COVID-19 transmission in Northeastern Brazil: Prospective space-time modeling. *Epidemiology and Infection*. https://doi.org/10.1017/S0950268820001843

Goutte, S., Péran, T., & Porcher, T. (2020). The role of economic structural factors in determining pandemic mortality rates: Evidence from the COVID-19 outbreak in France. *Research in International Business and Finance*. https://doi.org/10.1016/j.ribaf.2020.101281

Government of Canada. (2020). *Coronavirus disease 2019 (COVID-19): Epidemiology update*. https://health-infobase.canada.ca/covid-19/epidemiological-summary-covid-19-cases.html

Government of Canada. (2021a). *Prevalence of low income – National Household Survey (NHS) Dictionary*. https://www12.statcan.gc.ca/nhs-enm/2011/ref/dict/fam025-eng.cfm

Government of Canada, S. C. (2017a, February 8). *Focus on Geography Series, 2016 Census—Canada*. https://www12.statcan.gc.ca/census-recensement/2016/as-sa/fogs-spg/Facts-can-eng.cfm?Lang=Eng&GK=CAN&GC=01&TOPIC=7

Government of Canada, S. C. (2017b, May 3). *Dictionary, Census of Population, 2016—Highest certificate, diploma or degree*. https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/pop038-eng.cfm

Government of Canada, S. C. (2017c, May 3). *Dictionary, Census of Population, 2016—Immigrant*. https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/pop221-eng.cfm

Government of Canada, S. C. (2017d, May 3). *Dictionary, Census of Population, 2016—Low-income cut-offs, after tax (LICO-AT)*. https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/fam019-eng.cfm

Government of Canada, S. C. (2017e, May 3). *Dictionary, Census of Population, 2016—Prevalence of low income*. https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/fam025-eng.cfm

Government of Canada, S. C. (2017f, May 3). *Dictionary, Census of Population, 2016—Unemployment rate*. https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/pop125-eng.cfm

Government of Canada, S. C. (2017g, May 3). *Dictionary, Census of Population, 2016—Visible minority*. https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/pop127-eng.cfm

Government of Canada, S. C. (2017h, November 15). *Dictionary, Census of Population, 2016—Core housing need*. https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/households-menage037-eng.cfm

Government of Canada, S. C. (2021b, March 11). *COVID-19 in Canada: A One-year Update on Social and Economic Impacts*. https://www150.statcan.gc.ca/n1/pub/11-631-x/11-631-x2021001-eng.htm

Government of Canada, S. C. (2021c, April 9). *Labour force characteristics by age group, monthly, seasonally adjusted*. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1410028702

Graham, A. J., Atkinson, P. M., & Danson, F. M. (2004). Spatial analysis for epidemiology. *Acta Tropica*. https://doi.org/10.1016/j.actatropica.2004.05.001

Greene, S. K., Peterson, E. R., Balan, D., Jones, L., Culp, G. M., Fine, A. D., & Kulldorff, M. (2020). Detecting emerging COVID-19 community outbreaks at high spatiotemporal resolution—New York City, June–July 2020. In *MedRxiv*. https://doi.org/10.1101/2020.07.18.20156901

Grekousis, G., Feng, Z., Marakakis, I., Lu, Y., & Wang, R. (2022). Ranking the importance of demographic, socioeconomic, and underlying health factors on US COVID-19 deaths: A geographical random forest approach. *Health & Place*, *74*, 102744. https://doi.org/10.1016/j.healthplace.2022.102744

Griffith, D. A. (2010). Modeling spatio-temporal relationships: Retrospect and prospect. *Journal of Geographical Systems*, *12*(2), 111–123. https://doi.org/10.1007/s10109-010-0120-x

Grillet, M.-E., Barrera, R., Martínez, J.-E., Berti, J., & Fortin, M.-J. (2010). Disentangling the Effect of Local and Global Spatial Variation on a Mosquito-Borne Infection in a Neotropical Heterogeneous Environment. *The American Journal of Tropical Medicine and Hygiene*, *82*(2), 194–201. https://doi.org/10.4269/ajtmh.2010.09-0040

Gschlößl, S., & Czado, C. (2008). Modelling count data with overdispersion and spatial effects. *Statistical Papers*, *49*(3), 531–552. https://doi.org/10.1007/s00362-006-0031-6

Guan, W., Ni, Z., Hu, Y., Liang, W., Ou, C., He, J., Liu, L., Shan, H., Lei, C., Hui, D. S. C., Du, B., Li, L., Zeng, G., Yuen, K.-Y., Chen, R., Tang, C., Wang, T., Chen, P., Xiang, J., … Zhong, N. (2020). Clinical Characteristics of Coronavirus Disease 2019 in China. *New England Journal of Medicine*. https://doi.org/10.1056/nejmoa2002032

Haining, R., & Li, G. (2020). *Modelling Spatial and Spatial-Temporal Data: A Bayesian Approach* (1st ed.). CRC Press, Boca Raton, FL.

Han, Y., Yang, L., Jia, K., Li, J., Feng, S., Chen, W., Zhao, W., & Pereira, P. (2021). Spatial distribution characteristics of the COVID-19 pandemic in Beijing and its relationship with environmental factors. *Science of The Total Environment*, *761*, 144257. https://doi.org/10.1016/j.scitotenv.2020.144257

Harapan, H., Itoh, N., Yufika, A., Winardi, W., Keam, S., Te, H., Megawati, D., Hayati, Z., Wagner, A. L., & Mudatsir, M. (2020). Coronavirus disease 2019 (COVID-19): A literature review. *Journal of Infection and Public Health*, *13*(5), 667–673. https://doi.org/10.1016/j.jiph.2020.03.019

Hartley, D. M., & Perencevich, E. N. (2020). Public Health Interventions for COVID-19: Emerging Evidence and Implications for an Evolving Public Health Crisis. *JAMA*, *323*(19), 1908–1909. https://doi.org/10.1001/jama.2020.5910

Harwood, S., & Myers, D. (2002). The Dynamics of Immigration and Local Governance in Santa Ana. *Policy Studies Journal*, *30*(1), 70–91. https://doi.org/10.1111/j.1541-0072.2002.tb02130.x

Hasan Bhuiyan, M. T., Mahmud Khan, I., Rahman Jony, S. S., Robinson, R., Nguyen, U.-S. D. T., Keellings, D., Rahman, M. S., & Haque, U. (2021). The Disproportionate Impact of

COVID-19 among Undocumented Immigrants and Racial Minorities in the US. *International Journal of Environmental Research and Public Health*, *18*(23), Article 23. https://doi.org/10.3390/ijerph182312708

Hassan, Md. S., Bhuiyan, M. A. H., Tareq, F., Bodrud-Doza, Md., Tanu, S. M., & Rabbani, K. A. (2021). Relationship between COVID-19 infection rates and air pollution, geo-meteorological, and social parameters. *Environmental Monitoring and Assessment*, *193*(1), 29. https://doi.org/10.1007/s10661-020-08810-4

Hatchett, R. J., Mecher, C. E., & Lipsitch, M. (2007). Public health interventions and epidemic intensity during the 1918 influenza pandemic. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(18), 7582–7587. https://doi.org/10.1073/pnas.0610941104

Havulinna, A. S. (2011). *Bayesian spatial and temporal epidemiology of non-communicable diseases and mortality*. Aalto University. https://aaltodoc.aalto.fi:443/handle/123456789/5084

Hefley, T. J., Hooten, M. B., Hanks, E. M., Russell, R. E., & Walsh, D. P. (2017). Dynamic spatio-temporal models for spatial data. *Spatial Statistics*, *20*, 206–220. https://doi.org/10.1016/j.spasta.2017.02.005

Hempel, S., Burke, R. V., Hochman, M., Thompson, G., Brothers, A., Shin, J., Motala, A., Larkin, J., & Ringel, J. (2020). *Resource Allocation and Pandemic Response: An Evidence Synthesis To Inform Decision Making*. Agency for Healthcare Research and Quality (US). https://www.ncbi.nlm.nih.gov/books/NBK562921/

Hinne, M., Gronau, Q. F., van den Bergh, D., & Wagenmakers, E.-J. (2020). A Conceptual Introduction to Bayesian Model Averaging. *Advances in Methods and Practices in Psychological Science*, *3*(2), 200–215. https://doi.org/10.1177/2515245919898657

Hohl, A., Delmelle, E. M., Desjardins, M. R., & Lan, Y. (2020). Daily surveillance of COVID-19 using the prospective space-time scan statistic in the United States. *Spatial and Spatio-Temporal Epidemiology*. https://doi.org/10.1016/j.sste.2020.100354

Hohl, A., Delmelle, E., Tang, W., & Casas, I. (2016). Accelerating the discovery of space-time patterns of infectious diseases using parallel computing. *Spatial and Spatio-Temporal Epidemiology*, *19*, 10–20. https://doi.org/10.1016/j.sste.2016.05.002

Hu, T., Wang, S., She, B., Zhang, M., Huang, X., Cui, Y., Khuri, J., Hu, Y., Fu, X., Wang, X., Wang, P., Zhu, X., Bao, S., Guan, W., & Li, Z. (2021). Human mobility data in the COVID-19 pandemic: Characteristics, applications, and challenges. *International Journal of Digital Earth*, *14*(9), 1126–1147. https://doi.org/10.1080/17538947.2021.1952324

Hu, T., Yue, H., Wang, C., She, B., Ye, X., Liu, R., Zhu, X., Guan, W. W., & Bao, S. (2020). Racial Segregation, Testing Site Access, and COVID-19 Incidence Rate in Massachusetts, USA. *International Journal of Environmental Research and Public Health*, *17*(24), 9528. https://doi.org/10.3390/ijerph17249528

Huang, J., Kwan, M.-P., Kan, Z., Wong, M. S., Kwok, C. Y. T., & Yu, X. (2020). Investigating the Relationship between the Built Environment and Relative Risk of COVID-19 in Hong Kong. *ISPRS International Journal of Geo-Information*, *9*(11), 624. https://doi.org/10.3390/ijgi9110624

Huang, Z. (2021). Spatiotemporal Evolution Patterns of the COVID-19 Pandemic Using Space-Time Aggregation and Spatial Statistics: A Global Perspective. *ISPRS International Journal of Geo-Information*, *10*(8), Article 8. https://doi.org/10.3390/ijgi10080519

Hughes, G. J., & Gorton, R. (2013). An evaluation of SaTScan for the prospective detection of space-time Campylobacter clusters in the North East of England. *Epidemiology & Infection*, *141*(11), 2354–2364. https://doi.org/10.1017/S0950268812003135

Ibáñez-Beroiz, B., Librero-López, J., Peiró-Moreno, S., & Bernal-Delgado, E. (2011). Shared component modelling as an alternative to assess geographical variations in medical practice: Gender inequalities in hospital admissions for chronic diseases. *BMC Medical Research Methodology*, *11*(1), 172. https://doi.org/10.1186/1471-2288-11-172

Iftekhar, E. N., Priesemann, V., Balling, R., Bauer, S., Beutels, P., Calero Valdez, A., Cuschieri, S., Czypionka, T., Dumpis, U., Glaab, E., Grill, E., Hanson, C., Hotulainen, P., Klimek, P., Kretzschmar, M., Krüger, T., Krutzinna, J., Low, N., Machado, H., … Willeit, P. (2021). A look into the future of the COVID-19 pandemic in Europe: An expert consultation. *The Lancet Regional Health - Europe*, *8*, 100185. https://doi.org/10.1016/j.lanepe.2021.100185

Imai, N., Gaythorpe, K. A. M., Abbott, S., Bhatia, S., van Elsland, S., Prem, K., Liu, Y., & Ferguson, N. M. (2020). Adoption and impact of non-pharmaceutical interventions for COVID-19. *Wellcome Open Research*, *5*, 59. https://doi.org/10.12688/wellcomeopenres.15808.1

Indseth, T., Grøsland, M., Arnesen, T., Skyrud, K., Kløvstad, H., Lamprini, V., Telle, K., & Kjøllesdal, M. (2021). COVID-19 among immigrants in Norway, notified infections,

related hospitalizations and associated mortality: A register-based study. *Scandinavian Journal of Public Health*, *49*(1), 48–56. https://doi.org/10.1177/1403494820984026

Ingen, T. van, Akingbola, S., Brown, K. A., Daneman, N., Buchan, S. A., & Smith, B. T. (2021). *Neighbourhood-level risk factors of COVID-19 incidence and mortality* (p. 2021.01.27.21250618). https://doi.org/10.1101/2021.01.27.21250618

Islam, A., Sayeed, Md. A., Rahman, Md. K., Ferdous, J., Islam, S., & Hassan, M. M. (2021). Geospatial dynamics of COVID-19 clusters and hotspots in Bangladesh. *Transboundary and Emerging Diseases*. https://doi.org/10.1111/tbed.13973

Iyanda, A. E., Adeleke, R., Lu, Y., Osayomi, T., Adaralegbe, A., Lasode, M., Chima-Adaralegbe, N. J., & Osundina, A. M. (2020). A retrospective cross-national examination of COVID-19 outbreak in 175 countries: A multiscale geographically weighted regression analysis (January 11-June 28, 2020). *Journal of Infection and Public Health*. https://doi.org/10.1016/j.jiph.2020.07.006

Iyanda, A. E., & Osayomi, T. (2021). Is there a relationship between economic indicators and road fatalities in Texas? A multiscale geographically weighted regression analysis. *GeoJournal*, *86*(6), 2787–2807. https://doi.org/10.1007/s10708-020-10232-1

Jalilian, A., & Mateu, J. (2021). A hierarchical spatio-temporal model to analyze relative risk variations of COVID-19: A focus on Spain, Italy and Germany. *Stochastic Environmental Research and Risk Assessment*, *35*(4), 797–812. https://doi.org/10.1007/s00477-021-02003-2

Jaya, I. G. N. M., & Folmer, H. (2021). Bayesian spatiotemporal forecasting and mapping of COVID-19 risk with application to West Java Province, Indonesia. *Journal of Regional Science*, *61*(4), 849–881. https://doi.org/10.1111/jors.12533

Jeefoo, P., Tripathi, N. K., & Souris, M. (2011). Spatio-temporal diffusion pattern and hotspot detection of dengue in Chachoengsao province, Thailand. *International Journal of Environmental Research and Public Health*. https://doi.org/10.3390/ijerph8010051

Jerrett, M., Gale, S., & Kontgis, C. (2010). Spatial Modeling in Environmental and Public Health Research. *International Journal of Environmental Research and Public Health*, *7*(4), 1302–1329. https://doi.org/10.3390/ijerph7041302

Jia, P., Stein, A., James, P., Brownson, R. C., Wu, T., Xiao, Q., Wang, L., Sabel, C. E., & Wang, Y. (2019). Earth Observation: Investigating Noncommunicable Diseases from Space. *Annual Review of Public Health*, *40*(1), 85–104. https://doi.org/10.1146/annurev-publhealth-040218-043807

Jia, W.-J., Wang, M.-F., Zhou, C.-H., & Yang, Q.-H. (2021). Analysis of the spatial association of geographical detector-based landslides and environmental factors in the southeastern Tibetan Plateau, China. *PLoS ONE*, *16*(5), e0251776. https://doi.org/10.1371/journal.pone.0251776

Jiang, B. (2015). Geospatial analysis requires a different way of thinking: The problem of spatial heterogeneity. *GeoJournal*, *80*(1), 1–13. https://doi.org/10.1007/s10708-014-9537-y

Jiang, Y., Han, R., Su, J., Fan, X., Yu, H., Tao, R., & Zhou, J. (2022). Trends and predictions of lung cancer incidence in Jiangsu Province, China, 2009–2030: A bayesian age-period-

cohort modelling study. *BMC Cancer*, *22*(1), 1110. https://doi.org/10.1186/s12885-022-10187-1

Jiménez-Muñoz, J. C., Sobrino, J. A., Skoković, D., Mattar, C., & Cristóbal, J. (2014). Land Surface Temperature Retrieval Methods From Landsat-8 Thermal Infrared Sensor Data. *IEEE Geoscience and Remote Sensing Letters*, *11*(10), 1840–1843. https://doi.org/10.1109/LGRS.2014.2312032

Johnson, D. P., Ravi, N., & Braneon, C. V. (2021). Spatiotemporal Associations Between Social Vulnerability, Environmental Measurements, and COVID-19 in the Conterminous United States. *GeoHealth*, *5*(8), e2021GH000423. https://doi.org/10.1029/2021GH000423

Joinpoint Regression Program. (2011). https://surveillance.cancer.gov/joinpoint/

Jones, R. C. (2021). The spatial diffusion of covid-19 in Texas. *The Social Science Journal*, *0*(0), 1–14. https://doi.org/10.1080/03623319.2021.1926148

Jones, V. F., Lawson, P., Robson, G., Buchanan, B., & Aldrich, T. (2004). The Use of Spatial Statistics to Identify Asthma Risk Factors in an Urban Community. *Pediatric Asthma, Allergy & Immunology*, *17*(1), 3–13. https://doi.org/10.1089/088318704322994895

Kabisch, N., & Haase, D. (2014). Green justice or just green? Provision of urban green spaces in Berlin, Germany. *Landscape and Urban Planning*, *122*, 129–139. https://doi.org/10.1016/j.landurbplan.2013.11.016

Kadi, N., & Khelfaoui, M. (2020). Population density, a factor in the spread of COVID-19 in Algeria: Statistic study. *Bulletin of the National Research Centre*. https://doi.org/10.1186/s42269-020-00393-x

Kamel Boulos, M. N., & Geraghty, E. M. (2020). Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: How 21st century GIS technologies are supporting the global fight against outbreaks and epidemics. *International Journal of Health Geographics*, *19*(1), 8. https://doi.org/10.1186/s12942-020-00202-8

Kan, Z., Kwan, M.-P., Wong, M. S., Huang, J., & Liu, D. (2021). Identifying the space-time patterns of COVID-19 risk and their associations with different built environment features in Hong Kong. *The Science of the Total Environment*, *772*, 145379. https://doi.org/10.1016/j.scitotenv.2021.145379

Kang, D., Choi, H., Kim, J. H., & Choi, J. (2020). Spatial epidemic dynamics of the COVID-19 outbreak in China. *International Journal of Infectious Diseases*. https://doi.org/10.1016/j.ijid.2020.03.076

Karmarkar, E. N., Blanco, I., Amornkul, P. N., DuBois, A., Deng, X., Moonan, P. K., Rubenstein, B. L., Miller, D. A., Kennedy, I., Yu, J., Dauterman, J. P., Ongpin, M., Hathaway, W., Hoo, L., Trammell, S., Dosunmu, E. F., Yu, G., Khwaja, Z., Lu, W., … Williams, T. (2021). Timely intervention and control of a novel coronavirus (COVID-19) outbreak at a large skilled nursing facility—San Francisco, California, 2020. *Infection Control and Hospital Epidemiology*, 1–8. https://doi.org/10.1017/ice.2020.1375

Kathe, N., & Wani, R. (2020). Determinants of COVID-19 Incidence and Mortality in the US: Spatial Analysis. *MedRxiv*.

Keil, R., & Ali, H. (2007). Governing the Sick City: Urban Governance in the Age of Emerging Infectious Disease. *Antipode*, *39*(5), 846–873. https://doi.org/10.1111/j.1467-8330.2007.00555.x

Khanna, R. C., Cicinelli, M. V., Gilbert, S. S., Honavar, S. G., & Murthy, G. V. S. (2020). COVID-19 pandemic: Lessons learned and future directions. *Indian Journal of Ophthalmology*, *68*(5), 703–710. https://doi.org/10.4103/ijo.IJO_843_20

Khundi, M., Carpenter, J. R., Nliwasa, M., Cohen, T., Corbett, E. L., & MacPherson, P. (2021). Effectiveness of spatially targeted interventions for control of HIV, tuberculosis, leprosy and malaria: A systematic review. *BMJ Open*, *11*(7), e044715. https://doi.org/10.1136/bmjopen-2020-044715

Kim, B., Rundle, A. G., Goodwin, A. T. S., Morrison, C. N., Branas, C. C., El-Sadr, W., & Duncan, D. T. (2021). COVID-19 testing, case, and death rates and spatial socio-demographics in New York City: An ecological analysis as of June 2020. *Health & Place*, *68*, 102539. https://doi.org/10.1016/j.healthplace.2021.102539

Kim, H. J., Fay, M. P., Feuer, E. J., & Midthune, D. N. (2000). Permutation tests for joinpoint regression with applications to cancer rates. *Statistics in Medicine*, *19*(3), 335–351. https://doi.org/10.1002/(sici)1097-0258(20000215)19:3<335::aid-sim336>3.0.co;2-z

Kim, S., & Castro, M. C. (2020). Spatiotemporal pattern of COVID-19 and government response in South Korea (as of May 31, 2020). *International Journal of Infectious Diseases*. https://doi.org/10.1016/j.ijid.2020.07.004

Kindi, K. M. A., Al-Mawali, A., Akharusi, A., Alshukaili, D., Alnasiri, N., Al-Awadhi, T., Charabi, Y., & Kenawy, A. M. E. (2021). Demographic and socioeconomic determinants

of COVID-19 across Oman—A geospatial modelling approach. *Geospatial Health*, *16*(1), Article 1. https://doi.org/10.4081/gh.2021.985

Kirby, R. S., Delmelle, E., & Eberth, J. M. (2017). Advances in spatial epidemiology and geographic information systems. *Annals of Epidemiology*. https://doi.org/10.1016/j.annepidem.2016.12.001

Kivimäki, M., Batty, G. D., Pentti, J., Nyberg, S. T., Lindbohm, J. V., Ervasti, J., Gonzales-Inca, C., Suominen, S. B., Stenholm, S., Sipilä, P. N., Dadvand, P., & Vahtera, J. (2021). Modifications to residential neighbourhood characteristics and risk of 79 common health conditions: A prospective cohort study. *The Lancet Public Health*, *6*(6), e396–e407. https://doi.org/10.1016/S2468-2667(21)00066-9

Knorr-Held, L. (2000). Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine*, *19*(17–18), 2555–2567. https://doi.org/10.1002/1097-0258(20000915/30)19:17/18<2555::AID-SIM587>3.0.CO;2-#

Knorr-Held, L., & Besag, J. (1998). Modelling risk from a disease in time and space. *Statistics in Medicine*, *17*(18), 2045–2060. https://doi.org/10.1002/(SICI)1097-0258(19980930)17:18<2045::AID-SIM943>3.0.CO;2-P

Knorr-Held, L., & Best, N. G. (2001). A shared component model for detecting joint and selective clustering of two diseases. *Journal of the Royal Statistical Society. Series A: Statistics in Society*. https://doi.org/10.1111/1467-985X.00187

Knorr-Held, L., & Raßer, G. (2000). Bayesian detection of clusters and discontinuities in disease maps. *Biometrics*. https://doi.org/10.1111/j.0006-341X.2000.00013.x

Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., Open COVID-19 Data Working Group, du Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, *368*(6490), 493–497. https://doi.org/10.1126/science.abb4218

Kreis, C., Doessegger, E., Lupatsch, J. E., & Spycher, B. D. (2019). Space–time clustering of childhood cancers: A systematic review and pooled analysis. *European Journal of Epidemiology*, *34*(1), 9–21. https://doi.org/10.1007/s10654-018-0456-y

Kroll, M., Phalkey, R. K., & Kraas, F. (2015). Challenges to the surveillance of non-communicable diseases—A review of selected approaches. *BMC Public Health*. https://doi.org/10.1186/s12889-015-2570-z

Kronenberg, J., Haase, A., Łaszkiewicz, E., Antal, A., Baravikova, A., Biernacka, M., Dushkova, D., Filčak, R., Haase, D., Ignatieva, M., Khmara, Y., Niţă, M. R., & Onose, D. A. (2020). Environmental justice in the context of urban green space availability, accessibility, and attractiveness in postsocialist cities. *Cities*, *106*, 102862. https://doi.org/10.1016/j.cities.2020.102862

Kuebart, A., & Stabler, M. (2020). Infectious Diseases as Socio-Spatial Processes: The COVID-19 Outbreak In Germany. *Tijdschrift Voor Economische En Sociale Geografie = Journal of Economic and Social Geography = Revue De Geographie Economique Et Humaine = Zeitschrift Fur Okonomische Und Soziale Geographie = Revista De Geografia Economica Y Social*. https://doi.org/10.1111/tesg.12429

Kulldorff, M. (1999). Spatial Scan Statistics: Models, Calculations, and Applications. In J. Glaz & N. Balakrishnan (Eds.), *Scan Statistics and Applications* (pp. 303–322). Birkhäuser. https://doi.org/10.1007/978-1-4612-1578-3_14

Kulldorff, M. & Information Management Services Inc. (2009). SaTScan v9.3: Software for the spatial and space-time scan statistics. *StatScan, Boston, USA*.

Kulu, H., & Dorey, P. (2021). Infection rates from Covid-19 in Great Britain by geographical units: A model-based estimation from mortality data. *Health & Place*, *67*, 102460. https://doi.org/10.1016/j.healthplace.2020.102460

Ladoy, A., Opota, O., Carron, P.-N., Guessous, I., Vuilleumier, S., Joost, S., & Greub, G. (2021). Size and duration of COVID-19 clusters go along with a high SARS-CoV-2 viral load: A spatio-temporal investigation in Vaud state, Switzerland. *Science of The Total Environment*, *787*, 147483. https://doi.org/10.1016/j.scitotenv.2021.147483

Laohasiriwong, W., Puttanapong, N., & Luenam, A. (2018). A comparison of spatial heterogeneity with local cluster detection methods for chronic respiratory diseases in Thailand. *F1000Research*, *6*. https://doi.org/10.12688/f1000research.12128.2

Law, J., Haining, R., Maheswaran, R., & Pearson, T. (2006). Analyzing the Relationship Between Smoking and Coronary Heart Disease at the Small Area Level: A Bayesian Approach to Spatial Modeling. *Geographical Analysis*, *38*(2), 140–159. https://doi.org/10.1111/j.0016-7363.2006.00680.x

Law, J., Quick, M., & Chan, P. (2014). Bayesian Spatio-Temporal Modeling for Analysing Local Patterns of Crime Over Time at the Small-Area Level. *Journal of Quantitative Criminology*. https://doi.org/10.1007/s10940-013-9194-1

Law, J., Quick, M., & Chan, P. W. (2015). Analyzing hotspots of crime using a bayesian spatiotemporal modeling approach: A case study of violent crime in the greater Toronto area. *Geographical Analysis*. https://doi.org/10.1111/gean.12047

Lawson, A. B. (2010). Hotspot detection and clustering: Ways and means. *Environmental and Ecological Statistics*, *17*(2), 231–245. https://doi.org/10.1007/s10651-010-0142-z

Leal-Neto, O. B., Santos, F. A. S., Lee, J. Y., Albuquerque, J. O., & Souza, W. V. (2020). Prioritizing COVID-19 tests based on participatory surveillance and spatial scanning. *International Journal of Medical Informatics*. https://doi.org/10.1016/j.ijmedinf.2020.104263

Lee, D., & Sarran, C. (2015). Controlling for unmeasured confounding and spatial misalignment in long-term air pollution and health studies. *Environmetrics*, *26*(7), 477–487. https://doi.org/10.1002/env.2348

Lee, S. S., & Wong, N. S. (2011). The clustering and transmission dynamics of pandemic influenza A (H1N1) 2009 cases in Hong Kong. *Journal of Infection*, *63*(4), 274–280. https://doi.org/10.1016/j.jinf.2011.03.011

Leong, Y.-Y., & Yue, J. C. (2017). A modification to geographically weighted regression. *International Journal of Health Geographics*, *16*(1), 11. https://doi.org/10.1186/s12942-017-0085-9

Leung, Y., Mei, C.-L., & Zhang, W.-X. (2003). Statistical Test for Local Patterns of Spatial Association. *Environment and Planning A: Economy and Space*, *35*(4), 725–744. https://doi.org/10.1068/a3550

Li, B., Peng, Y., He, H., Wang, M., & Feng, T. (2021). Built environment and early infection of COVID-19 in urban districts: A case study of Huangzhou. *Sustainable Cities and Society*, *66*, 102685. https://doi.org/10.1016/j.scs.2020.102685

Li, C.-S., Lu, J.-C., Park, J., Kim, K., Brinkley, P. A., & Peterson, J. P. (1999). Multivariate Zero-Inflated Poisson Models and Their Applications. *Technometrics*, *41*(1), 29–38. https://doi.org/10.1080/00401706.1999.10485593

Li, G., Best, N., Hansell, A. L., Ahmed, I., & Richardson, S. (2012). BaySTDetect: Detecting unusual temporal patterns in small area data via Bayesian model choice. *Biostatistics (Oxford, England)*, *13*(4), 695–710. https://doi.org/10.1093/biostatistics/kxs005

Li, H., Li, H., Ding, Z., Hu, Z., Chen, F., Wang, K., Peng, Z., & Shen, H. (2020). Spatial statistical analysis of coronavirus disease 2019 (Covid-19) in China. *Geospatial Health*. https://doi.org/10.4081/gh.2020.867

Li, H., Wang, S., Zhong, F., Bao, W., Li, Y., Liu, L., Wang, H., & He, Y. (2020). Age-Dependent Risks of Incidence and Mortality of COVID-19 in Hubei Province and Other Parts of China. *Frontiers in Medicine*. https://doi.org/10.3389/fmed.2020.00190

Li, M., Ma, S., & Liu, Z. (2022). A novel method to detect the early warning signal of COVID-19 transmission. *BMC Infectious Diseases*, *22*(1), 626. https://doi.org/10.1186/s12879-022-07603-z

Li, W., Gong, J., Zhou, J., Fan, H., Qin, C., Gong, Y., & Hu, W. (2022). The Analysis of Patterns of Two COVID-19 Outbreak Clusters in China. *International Journal of Environmental Research and Public Health*, *19*(8), Article 8. https://doi.org/10.3390/ijerph19084876

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., & Sobrino, J. A. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment, 131,* 14–37. https://doi.org/10.1016/j.rse.2012.12.008

Liao, J., & Brookmeyer, R. (1995). An Empirical Bayes Approach to Smoothing in Backcalculation of HIV Infection Rates. *Biometrics*, *51*(2), 579–588. https://doi.org/10.2307/2532946

Lima, E. E. C. de, Gayawan, E., Baptista, E. A., & Queiroz, B. L. (2021). Spatial pattern of COVID-19 deaths and infections in small areas of Brazil. *PloS One*, *16*(2), e0246808. https://doi.org/10.1371/journal.pone.0246808

Lin, C. H., & Wen, T. H. (2011). Using geographically weighted regression (GWR) to explorespatial varying relationships of immature mosquitoes andhuman densities with the incidence of dengue. *International Journal of Environmental Research and Public Health*. https://doi.org/10.3390/ijerph8072798

Lin, C.-H., & Wen, T.-H. (2022). How Spatial Epidemiology Helps Understand Infectious Human Disease Transmission. *Tropical Medicine and Infectious Disease*, *7*(8), Article 8. https://doi.org/10.3390/tropicalmed7080164

Lin, Y., Zhong, P., & Chen, T. (2020). Association Between Socioeconomic Factors and the COVID-19 Outbreak in the 39 Well-Developed Cities of China. *Frontiers in Public Health*. https://doi.org/10.3389/fpubh.2020.546637

Liu, F., Wang, J., Liu, J., Li, Y., Liu, D., Tong, J., Li, Z., Yu, D., Fan, Y., Bi, X., Zhang, X., & Mo, S. (2020). Predicting and analyzing the COVID-19 epidemic in China: Based on

SEIRD, LSTM and GWR models. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0238280

Liu, L., Hu, T., Bao, S., Wu, H., Peng, Z., & Wang, R. (2021). The Spatiotemporal Interaction Effect of COVID-19 Transmission in the United States. *ISPRS International Journal of Geo-Information*, *10*(6), Article 6. https://doi.org/10.3390/ijgi10060387

Lo, C.-H., Nguyen, L. H., Drew, D. A., Warner, E. T., Joshi, A. D., Graham, M. S., Anyane-Yeboa, A., Shebl, F. M., Astley, C. M., Figueiredo, J. C., Guo, C.-G., Ma, W., Mehta, R. S., Kwon, S., Song, M., Davies, R., Capdevila, J., Sudre, C. H., Wolf, J., … Chan, A. T. (2021). Race, ethnicity, community-level socioeconomic factors, and risk of COVID-19 in the United States and the United Kingdom. *EClinicalMedicine*, *38*. https://doi.org/10.1016/j.eclinm.2021.101029

Lombardo, J. S., & Buckeridge, D. L. (2006). Disease Surveillance: A Public Health Informatics Approach. In *Disease Surveillance: A Public Health Informatics Approach*. https://doi.org/10.1002/9780470131886

Long, J. A., Malekzadeh, M., Klar, B., & Martin, G. (2021). Do regionally targeted lockdowns alter movement to non-lockdown regions? Evidence from Ontario, Canada. *Health & Place*, 102668. https://doi.org/10.1016/j.healthplace.2021.102668

Long, Y., Chen, Y., & Li, Y. (2020). Multifractal scaling analyses of the spatial diffusion pattern of COVID-19 pandemic in Chinese mainland. *Cornell University, Physics an*.

López-Quílez, A. (2019). Spatio-Temporal Analysis of Infectious Diseases. *International Journal of Environmental Research and Public Health*, *16*(4), Article 4. https://doi.org/10.3390/ijerph16040669

Lowe, R., Bailey, T. C., Stephenson, D. B., Graham, R. J., Coelho, C. A. S., Sá Carvalho, M., & Barcellos, C. (2011). Spatio-temporal modelling of climate-sensitive disease risk: Towards an early warning system for dengue in Brazil. *Computers and Geosciences*. https://doi.org/10.1016/j.cageo.2010.01.008

Lowe, R., Chirombo, J., & Tompkins, A. M. (2013). Relative importance of climatic, geographic and socio-economic determinants of malaria in Malawi. *Malaria Journal*, *12*(1), 416. https://doi.org/10.1186/1475-2875-12-416

Ma, J., Zhu, H., Li, P., Liu, C., Li, F., Luo, Z., Zhang, M., & Li, L. (2022). Spatial Patterns of the Spread of COVID-19 in Singapore and the Influencing Factors. *ISPRS International Journal of Geo-Information*, *11*(3), Article 3. https://doi.org/10.3390/ijgi11030152

Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., Yan, J., Niu, J., Zhou, J., & Luo, B. (2020). Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.138226

MacNab, Y. C. (2022). Bayesian disease mapping: Past, present, and future. *Spatial Statistics*, *50*, 100593. https://doi.org/10.1016/j.spasta.2022.100593

MacNab, Y. C., Farrell, P. J., Gustafson, P., & Wen, S. (2004). Estimation in Bayesian Disease Mapping. *Biometrics*, *60*(4), 865–873. https://doi.org/10.1111/j.0006-341X.2004.00241.x

MacNab, Y. C., & Gustafson, P. (2007). Regression B-spline smoothing in Bayesian disease mapping: With an application to patient safety surveillance. *Statistics in Medicine*, *26*(24), 4455–4474. https://doi.org/10.1002/sim.2868

Madhav, N., Oppenheim, B., Gallivan, M., Mulembakani, P., Rubin, E., & Wolfe, N. (2017). Pandemics: Risks, Impacts, and Mitigation. In *Disease Control Priorities: Improving*

Health and Reducing Poverty. 3rd edition. The International Bank for Reconstruction and Development / The World Bank. https://doi.org/10.1596/978-1-4648-0527-1_ch17

Maiti, A., Zhang, Q., Sannigrahi, S., Pramanik, S., Chakraborti, S., Cerda, A., & Pilla, F. (2021). Exploring spatiotemporal effects of the driving factors on COVID-19 incidences in the contiguous United States. *Sustainable Cities and Society*, *68*, 102784. https://doi.org/10.1016/j.scs.2021.102784

Maiti, A., Zhang, Q., Sannigrahi, S., Pramanik, S., Chakraborti, S., & Pilla, F. (2020). Spatiotemporal effects of the causal factors on COVID-19 incidences in the contiguous United States. *ArXiv Preprint ArXiv:2010.15754*.

Manley, D. (2014). Scale, Aggregation, and the Modifiable Areal Unit Problem. In M. M. Fischer & P. Nijkamp (Eds.), *Handbook of Regional Science* (pp. 1157–1171). Springer. https://doi.org/10.1007/978-3-642-23430-9_69

Mansour, S., Al Kindi, A., Al-Said, A., Al-Said, A., & Atkinson, P. (2021). Sociodemographic determinants of COVID-19 incidence rates in Oman: Geospatial modelling using multiscale geographically weighted regression (MGWR). *Sustainable Cities and Society*, *65*, 102627. https://doi.org/10.1016/j.scs.2020.102627

Marchand-Senécal, X., Kozak, R., Mubareka, S., Salt, N., Gubbay, J. B., Eshaghi, A., Allen, V., Li, Y., Bastien, N., Gilmour, M., Ozaldin, O., & Leis, J. A. (2020). Diagnosis and Management of First Case of COVID-19 in Canada: Lessons Applied from SARS-CoV-1. *Clinical Infectious Diseases*. https://doi.org/10.1093/cid/ciaa227

Maritz, J. S., & Lwin, T. (Eds.). (2017). *Empirical Bayes Methods* (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781351071666

Markel, H., Lipman, H. B., Navarro, J. A., Sloan, A., Michalsen, J. R., Stern, A. M., & Cetron, M. S. (2007). Nonpharmaceutical Interventions Implemented by US Cities During the 1918-1919 Influenza Pandemic. *JAMA*, *298*(6), 644–654. https://doi.org/10.1001/jama.298.6.644

Maroko, A. R., Nash, D., & Pavilonis, B. T. (2020). COVID-19 and Inequity: A Comparative Spatial Analysis of New York City and Chicago Hot Spots. *Journal of Urban Health : Bulletin of the New York Academy of Medicine*, *97*(4), 461–470. https://doi.org/10.1007/s11524-020-00468-0

Marshall, R. J. (1991). A Review of Methods for the Statistical Analysis of Spatial Patterns of Disease. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *154*(3), 421–441. https://doi.org/10.2307/2983152

Mason, C. H., & Perreault, W. D. (1991). Collinearity, Power, and Interpretation of Multiple Regression Analysis. *Journal of Marketing Research*. https://doi.org/10.2307/3172863

Masrur, A., Yu, M., Luo, W., & Dewan, A. (2020). Space-time patterns, change, and propagation of covid-19 risk relative to the intervention scenarios in bangladesh. *International Journal of Environmental Research and Public Health*. https://doi.org/10.3390/ijerph17165911

Mathieu, E., Ritchie, H., Ortiz-Ospina, E., Roser, M., Hasell, J., Appel, C., Giattino, C., & Rodés-Guirao, L. (2021). A global database of COVID-19 vaccinations. *Nature Human Behaviour*, *5*(7), Article 7. https://doi.org/10.1038/s41562-021-01122-8

Matthew, R. A., & McDonald, B. (2006). Cities under Siege: Urban Planning and the Threat of Infectious Disease. *Journal of the American Planning Association*, *72*(1), 109–117. https://doi.org/10.1080/01944360608976728

Maude, R. J., Nguon, C., Ly, P., Bunkea, T., Ngor, P., Canavati de la Torre, S. E., White, N. J., Dondorp, A. M., Day, N. P., White, L. J., & Chuor, C. M. (2014). Spatial and temporal epidemiology of clinical malaria in Cambodia 2004–2013. *Malaria Journal*, *13*(1), 385. https://doi.org/10.1186/1475-2875-13-385

McCall, S. J., El Khoury, T., Salibi, N., Abi Zeid, B., El Haddad, M., Alawieh, M. F., Abdulrahim, S., Chaaya, M., Ghattas, H., & Sibai, A. M. (2022). Development of a Prediction Model for the Management of Noncommunicable Diseases Among Older Syrian Refugees Amidst the COVID-19 Pandemic in Lebanon. *JAMA Network Open*, *5*(10), e2231633. https://doi.org/10.1001/jamanetworkopen.2022.31633

McCandless, L. C., & Gustafson, P. (2017). A comparison of Bayesian and Monte Carlo sensitivity analysis for unmeasured confounding. *Statistics in Medicine*, *36*(18), 2887–2901. https://doi.org/10.1002/sim.7298

Mei, C. (2004). *Geographically Weighted Regression Technique for Spatial Data Analysis*.

Mena, C., Sepúlveda, C., Fuentes, E., Ormazábal, Y., & Palomo, I. (2018). Spatial analysis for the epidemiological study of cardiovascular diseases: A systematic literature search. *Geospatial Health*, *13*(1), Article 1. https://doi.org/10.4081/gh.2018.587

Merow, C., & Urban, M. C. (2020). Seasonality and uncertainty in global COVID-19 growth rates. *Proceedings of the National Academy of Sciences*, *117*(44), 27456–27464. https://doi.org/10.1073/pnas.2008590117

Meyer, R. (2016). Deviance Information Criterion (DIC). In *Wiley StatsRef: Statistics Reference Online* (pp. 1–6). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118445112.stat07878

Middya, A. I., & Roy, S. (2021). Geographically varying relationships of COVID-19 mortality with different factors in India. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-86987-5

Midi, H., Sarkar, S. K., & Rana, S. (2010). Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics*, *13*(3), 253–267. https://doi.org/10.1080/09720502.2010.10700699

Miller, A. K., Gordon, J. C., Curtis, J. W., Ajayakumar, J., Schumacher, F. R., & Avril, S. (2022). The Geographic Context of Racial Disparities in Aggressive Endometrial Cancer Subtypes: Integrating Social and Environmental Aspects to Discern Biological Outcomes. *International Journal of Environmental Research and Public Health*, *19*(14), Article 14. https://doi.org/10.3390/ijerph19148613

Miller, H. J. (2004). Tobler's First Law and Spatial Analysis. *Annals of the Association of American Geographers*, *94*(2), 284–289. https://doi.org/10.1111/j.1467-8306.2004.09402005.x

Millett, G. A., Jones, A. T., Benkeser, D., Baral, S., Mercer, L., Beyrer, C., Honermann, B., Lankiewicz, E., Mena, L., Crowley, J. S., Sherwood, J., & Sullivan, P. S. (2020). Assessing differential impacts of COVID-19 on black communities. *Annals of Epidemiology*, *47*, 37–44. https://doi.org/10.1016/j.annepidem.2020.05.003

Milman, O., Yelin, I., Aharony, N., Katz, R., Herzel, E., Ben-Tov, A., Kuint, J., Gazit, S., Chodick, G., Patalon, T., & Kishony, R. (2021). Community-level evidence for SARS-CoV-2 vaccine protection of unvaccinated individuals. *Nature Medicine*, *27*(8), Article 8. https://doi.org/10.1038/s41591-021-01407-5

Mohammadinia, A., Saeidian, B., Pradhan, B., & Ghaemi, Z. (2019). Prediction mapping of human leptospirosis using ANN, GWR, SVM and GLM approaches. *BMC Infectious Diseases*, *19*(1), 971. https://doi.org/10.1186/s12879-019-4580-4

Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.138884

Mondal, P., Sinharoy, A., & Su, L. (2021). Sociodemographic predictors of COVID-19 vaccine acceptance: A nationwide US-based survey study. *Public Health*, *198*, 252–259. https://doi.org/10.1016/j.puhe.2021.07.028

Moraga, P., & Kulldorff, M. (2016). Detection of spatial variations in temporal trends with a quadratic function. *Statistical Methods in Medical Research*, *25*(4), 1422–1437. https://doi.org/10.1177/0962280213485312

Morris, M., Wheeler-Martin, K., Simpson, D., Mooney, S. J., Gelman, A., & DiMaggio, C. (2019). Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in stan. *Spatial and Spatio-Temporal Epidemiology*, *31*, 100301. https://doi.org/10.1016/j.sste.2019.100301

Moss, T., & Kar, B. (2020). SOCIO-ECONOMIC VULNERABILITY TO URBAN HEAT IN PHOENIX, ARIZONA AND DALLAS, TEXAS DURING JUNE 2020. *ISPRS Annals of*

the Photogrammetry, Remote Sensing and Spatial Information Sciences, *VI-3-W1-2020*, 59–66. https://doi.org/10.5194/isprs-annals-VI-3-W1-2020-59-2020

Murray, J., & Cohen, A. L. (2017). Infectious Disease Surveillance. *International Encyclopedia of Public Health*, 222–229. https://doi.org/10.1016/B978-0-12-803678-5.00517-8

Mwaba, J., Debes, A. K., Shea, P., Mukonka, V., Chewe, O., Chisenga, C., Simuyandi, M., Kwenda, G., Sack, D., Chilengi, R., & Ali, M. (2020). Identification of cholera hotspots in Zambia: A spatiotemporal analysis of cholera data from 2008 to 2017. *PLOS Neglected Tropical Diseases*, *14*(4), e0008227. https://doi.org/10.1371/journal.pntd.0008227

Nagy, M., & Negru, D. (2014). USING CLUSTERING SOFTWARE FOR EXPLORING SPATIAL AND TEMPORAL PATTERNS IN NON-COMMUNICABLE DISEASES. *European Scientific Journal, ESJ*, *10*(33), Article 33. https://doi.org/10.19044/esj.2014.v10n33p%p

Nakaya, T. (2016). *GWR4.09 User Manual*. https://raw.githubusercontent.com/gwrtools/gwr4/master/GWR4manual_409.pdf

NASA. (2021). *Atmospheric Correction Parameter Calculator*. https://atmcorr.gsfc.nasa.gov/

Nasiri, R., Akbarpour, S., Zali, A. R., Khodakarami, N., Boochani, M. H., Noory, A. R., & Soori, H. (2021). Spatio-temporal analysis of COVID-19 incidence rate using GIS: A case study-Tehran metropolitan, Iran. *GeoJournal*, 1–15. https://doi.org/10.1007/s10708-021-10438-x

Nazia, N., Ali, M., Jakariya, M., Nahar, Q., Yunus, M., & Emch, M. (2018). Spatial and population drivers of persistent cholera transmission in rural Bangladesh: Implications for

vaccine and intervention targeting. *Spatial and Spatio-Temporal Epidemiology*, *24*. https://doi.org/10.1016/j.sste.2017.09.001

Nazia, N., Butt, Z. A., Bedard, M. L., Tang, W.-C., Sehar, H., & Law, J. (2022). Methods Used in the Spatial and Spatiotemporal Analysis of COVID-19 Epidemiology: A Systematic Review. *International Journal of Environmental Research and Public Health*, *19*(14), Article 14. https://doi.org/10.3390/ijerph19148267

Neelon, B., Ghosh, P., & Loebs, P. F. (2013). A spatial Poisson hurdle model for exploring geographic variation in emergency department visits. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *176*(2), 389–413. https://doi.org/10.1111/j.1467-985X.2012.01039.x

Neiderud, C.-J. (2015). How urbanization affects the epidemiology of emerging infectious diseases. *Infection Ecology & Epidemiology*, *5*, 10.3402/iee.v5.27060. https://doi.org/10.3402/iee.v5.27060

Neinavaz, E., Skidmore, A. K., & Darvishzadeh, R. (2020). Effects of prediction accuracy of the proportion of vegetation cover on land surface emissivity and temperature using the NDVI threshold method. *International Journal of Applied Earth Observation and Geoinformation*, *85*, 101984. https://doi.org/10.1016/j.jag.2019.101984

Nekorchuk, D. M., Gebrehiwot, T., Lake, M., Awoke, W., Mihretie, A., & Wimberly, M. C. (2021). Comparing malaria early detection methods in a declining transmission setting in northwestern Ethiopia. *BMC Public Health*, *21*(1), 788. https://doi.org/10.1186/s12889-021-10850-5

News ·, C. B. C. (2020, May 12). *Lower income people, new immigrants at higher COVID-19 risk in Toronto, data suggests | CBC News*. CBC. https://www.cbc.ca/news/canada/toronto/low-income-immigrants-covid-19-infection-1.5566384

Ngwa, M. C., Ihekweazu, C., Okwor, T., Yennan, S., Williams, N., Elimian, K., Karaye, N. Y., Bello, I. W., & Sack, D. A. (2021). The cholera risk assessment in Kano State, Nigeria: A historical review, mapping of hotspots and evaluation of contextual factors. *PLOS Neglected Tropical Diseases*, *15*(1), e0009046. https://doi.org/10.1371/journal.pntd.0009046

Ngwira, A., Kumwenda, F., Munthali, E. C. S., & Nkolokosa, D. (2021). Spatial temporal distribution of COVID-19 risk during the early phase of the pandemic in Malawi. *PeerJ*, *9*, e11003. https://doi.org/10.7717/peerj.11003

Niedzwiedz, C. L., O'Donnell, C. A., Jani, B. D., Demou, E., Ho, F. K., Celis-Morales, C., Nicholl, B. I., Mair, F. S., Welsh, P., Sattar, N., Pell, J. P., & Katikireddi, S. V. (2020). Ethnic and socioeconomic differences in SARS-CoV-2 infection: Prospective cohort study using UK Biobank. *BMC Medicine*. https://doi.org/10.1186/s12916-020-01640-8

Nii-Trebi, N. I. (2017). Emerging and Neglected Infectious Diseases: Insights, Advances, and Challenges. *BioMed Research International*, *2017*, e5245021. https://doi.org/10.1155/2017/5245021

Niu, X., Yue, Y., Zhou, X., & Zhang, X. (2020). How Urban Factors Affect the Spatiotemporal Distribution of Infectious Diseases in Addition to Intercity Population Movement in

China. *ISPRS International Journal of Geo-Information*, *9*(11), Article 11. https://doi.org/10.3390/ijgi9110615

Odusanya, O. O., Odugbemi, B. A., Odugbemi, T. O., & Ajisegiri, W. S. (2020). COVID-19: A review of the effectiveness of non-pharmacological interventions. *The Nigerian Postgraduate Medical Journal*, *27*(4), 261–267. https://doi.org/10.4103/npmj.npmj_208_20

Olmo, J., & Sanso-Navarro, M. (2021). Modeling the spread of COVID-19 in New York City. *Papers in Regional Science: The Journal of the Regional Science Association International*. https://doi.org/10.1111/pirs.12615

Oluyomi, A. O., Gunter, S. M., Leining, L. M., Murray, K. O., & Amos, C. (2021). COVID-19 Community Incidence and Associated Neighborhood-Level Characteristics in Houston, Texas, USA. *International Journal of Environmental Research and Public Health*, *18*(4), 1495. https://doi.org/10.3390/ijerph18041495

Onozuka, D., & Hagihara, A. (2007). Geographic prediction of tuberculosis clusters in Fukuoka, Japan, using the space-time scan statistic. *BMC Infectious Diseases*, *7*(1), 26. https://doi.org/10.1186/1471-2334-7-26

Ontario. (2022). *COVID-19 testing and treatment*. Ontario.Ca. http://www.ontario.ca/page/covid-19-testing-and-treatment

Osei, F. B., Duker, A. A., & Stein, A. (2011). Hierarchical Bayesian modeling of the space-time diffusion patterns of cholera epidemic in Kumasi, Ghana. *Statistica Neerlandica*, *65*(1), 84–100. https://doi.org/10.1111/j.1467-9574.2010.00475.x

Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, A. S. (2019). mgwr: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. *ISPRS International Journal of Geo-Information*, *8*(6), Article 6. https://doi.org/10.3390/ijgi8060269

Oshan, T. M., Smith, J. P., & Fotheringham, A. S. (2020). Targeting the spatial context of obesity determinants via multiscale geographically weighted regression. *International Journal of Health Geographics*, *19*(1), 11. https://doi.org/10.1186/s12942-020-00204-6

Oshan, T., Wolf, L. J., Fotheringham, A. S., Kang, W., Li, Z., & Yu, H. (2019). A comment on geographically weighted regression with parameter-specific distance metrics. *International Journal of Geographical Information Science*, *33*(7), 1289–1299. https://doi.org/10.1080/13658816.2019.1572895

Paciorek, C. J. (2010). The importance of scale for spatial-confounding bias and precision of spatial regression estimators. *Statistical Science : A Review Journal of the Institute of Mathematical Statistics*, *25*(1), 107–125. https://doi.org/10.1214/10-STS326

Paul, R., Arif, A. A., Adeyemi, O., Ghosh, S., & Han, D. (2020). Progression of COVID-19 From Urban to Rural Areas in the United States: A Spatiotemporal Analysis of Prevalence Rates. *Journal of Rural Health*. https://doi.org/10.1111/jrh.12486

Payne, S. (2017). Chapter 5—Virus Transmission and Epidemiology. In S. Payne (Ed.), *Viruses* (pp. 53–60). Academic Press. https://doi.org/10.1016/B978-0-12-803109-4.00005-2

Pei, S., Kandula, S., & Shaman, J. (2020). Differential Effects of Intervention Timing on COVID-19 Spread in the United States. *MedRxiv*, 2020.05.15.20103655. https://doi.org/10.1101/2020.05.15.20103655

Pellis, L., Scarabel, F., Stage, H. B., Overton, C. E., Chappell, L. H. K., Lythgoe, K. A., Fearon, E., Bennett, E., Curran-Sebastian, J., Das, R., Fyles, M., Lewkowicz, H., Pang, X., Vekaria, B., Webb, L., House, T., & Hall, I. (2020). Challenges in control of Covid-19: Short doubling time and long delay to effect of interventions. *ArXiv:2004.00117 [q-Bio]*. http://arxiv.org/abs/2004.00117

Pequeno, P., Mendel, B., Rosa, C., Bosholn, M., Souza, J. L., Baccaro, F., Barbosa, R., & Magnusson, W. (2020). Air transportation, population density and temperature predict the spread of COVID-19 in Brazil. *PeerJ*, *8*, e9322. https://doi.org/10.7717/peerj.9322

Perles, M.-J., Sortino, J. F., & Mérida, M. F. (2021). The Neighborhood Contagion Focus as a Spatial Unit for Diagnosis and Epidemiological Action against COVID-19 Contagion in Urban Spaces: A Methodological Proposal for Its Detection and Delimitation. *International Journal of Environmental Research and Public Health*, *18*(6), Article 6. https://doi.org/10.3390/ijerph18063145

Politi, E., Lüders, A., Sankaran, S., Anderson, J., Van Assche, J., Spiritus-Beerden, E., Roblain, A., Phalet, K., Derluyn, I., Verelst, A., & Green, E. G. T. (2021). The impact of COVID-19 on the majority population, ethno-racial minorities, and immigrants: A systematic literature review on threat appraisals from an inter-group perspective. *European Psychologist*, *26*(4), 298–309. https://doi.org/10.1027/1016-9040/a000460

Potluri, R., Kumar, A., Maheshwari, V., Smith, C., Mathieu, V. O., Luhn, K., Callendret, B., & Bhandari, H. (2020). Impact of prophylactic vaccination strategies on Ebola virus transmission: A modeling analysis. *PLOS ONE*, *15*(4), e0230406. https://doi.org/10.1371/journal.pone.0230406

Pourghasemi, H. R., Pouyan, S., Heidari, B., Farajzadeh, Z., Fallah Shamsi, S. R., Babaei, S., Khosravi, R., Etemadi, M., Ghanbarian, G., Farhadi, A., Safaeian, R., Heidari, Z., Tarazkar, M. H., Tiefenbacher, J. P., Azmi, A., & Sadeghian, F. (2020). Spatial modeling, risk mapping, change detection, and outbreak trend analysis of coronavirus (COVID-19) in Iran (days between February 19 and June 14, 2020). *International Journal of Infectious Diseases*. https://doi.org/10.1016/j.ijid.2020.06.058

Pranzo, A. M. R., Dai Prà, E., & Besana, A. (2022). Epidemiological geography at work: An exploratory review about the overall findings of spatial analysis applied to the study of CoViD-19 propagation along the first pandemic year. *GeoJournal*. https://doi.org/10.1007/s10708-022-10601-y

Preston, V., Murdie, R., Wedlock, J., Agrawal, S., Anucha, U., D'addario, S., Kwak, M. J., Logan, J., & Murnaghan, A. M. (2009). Immigrants and homelessness—At risk in Canada's outer suburbs. *The Canadian Geographer / Le Géographe Canadien*, *53*(3), 288–304. https://doi.org/10.1111/j.1541-0064.2009.00264.x

Qi, C., Zhu, Y. C., Li, C. Y., Hu, Y. C., Liu, L. L., Zhang, D. D., Wang, X., She, K. L., Jia, Y., Liu, T. X., & Li, X. J. (2020). Epidemiological characteristics and spatial-temporal analysis of COVID-19 in Shandong Province, China. *Epidemiology and Infection*. https://doi.org/10.1017/S095026882000151X

Qi, H., Xiao, S., Shi, R., Ward, M. P., Chen, Y., Tu, W., Su, Q., Wang, W., Wang, X., & Zhang, Z. (2020). COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.138778

Quilty, B. J., Clifford, S., Hellewell, J., Russell, T. W., Kucharski, A. J., Flasche, S., Edmunds, W. J., Atkins, K. E., Foss, A. M., Waterlow, N. R., Abbas, K., Lowe, R., Pearson, C. A. B., Funk, S., Rosello, A., Knight, G. M., Bosse, N. I., Procter, S. R., Gore-Langton, G. R., … Davies, N. G. (2021). Quarantine and testing strategies in contact tracing for SARS-CoV-2: A modelling study. *The Lancet Public Health*, *6*(3), e175–e183. https://doi.org/10.1016/S2468-2667(20)30308-X

Quinn, S. C., & Kumar, S. (2014). Health Inequalities and Infectious Disease Epidemics: A Challenge for Global Health Security. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, *12*(5), 263–273. https://doi.org/10.1089/bsp.2014.0032

Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian Model Averaging for Linear Regression Models. *Journal of the American Statistical Association*, *92*(437), 179–191. https://doi.org/10.1080/01621459.1997.10473615

Ramírez, I. J., & Lee, J. (2020). COVID-19 Emergence and Social and Health Determinants in Colorado: A Rapid Spatial Analysis. *International Journal of Environmental Research and Public Health*, *17*(11), 3856. https://doi.org/10.3390/ijerph17113856

Ramírez-Aldana, R., Gomez-Verjan, J. C., & Bello-Chavolla, O. Y. (2020). Spatial analysis of COVID-19 spread in Iran: Insights into geographical and structural transmission determinants at a province level. *PLOS Neglected Tropical Diseases*, *14*(11), e0008875. https://doi.org/10.1371/journal.pntd.0008875

Rao, H., Shi, X., & Zhang, X. (2017). Using the kulldorff's scan statistical analysis to detect spatio-temporal clusters of tuberculosis in Qinghai Province, China, 2009-2016. *BMC Infectious Diseases*. https://doi.org/10.1186/s12879-017-2643-y

Rashed, E. A., Kodera, S., Gomez-Tames, J., & Hirata, A. (2020). Influence of absolute humidity, temperature and population density on COVID-19 spread and decay durations: Multi-prefecture study in Japan. *International Journal of Environmental Research and Public Health*. https://doi.org/10.3390/ijerph17155354

Rawat, S., & Deb, S. (2021). A spatio-temporal statistical model to analyze COVID-19 spread in the USA. *Journal of Applied Statistics*, *0*(0), 1–20. https://doi.org/10.1080/02664763.2021.1970122

Raymundo, C. E., Oliveira, M. C., Eleuterio, T. de A., André, S. R., Silva, M. G. da, Queiroz, E. R. da S., & Medronho, R. de A. (2021). Spatial analysis of COVID-19 incidence and the sociodemographic context in Brazil. *PLOS ONE*, *16*(3), e0247794. https://doi.org/10.1371/journal.pone.0247794

Reuters, Martell, A., & Warburton, M. (2021, April 1). Ontario 'pulling the emergency brake' with third COVID-19 lockdown as cases rise, ICU beds fill. *Reuters*. https://www.reuters.com/business/healthcare-pharmaceuticals/new-covid-cases-ontario-will-surge-without-stay-home-order-ontario-modeling-2021-04-01/

Reyes, O., Lee, E. C., Sah, P., Viboud, C., Chandra, S., & Bansal, S. (2018). Spatiotemporal Patterns and Diffusion of the 1918 Influenza Pandemic in British India. *American Journal of Epidemiology*. https://doi.org/10.1093/aje/kwy209

Richmond, H. L., Tome, J., Rochani, H., Fung, I. C. H., Shah, G. H., & Schwind, J. S. (2020). The use of penalized regression analysis to identify county-level demographic and socioeconomic variables predictive of increased covid-19 cumulative case rates in the

state of georgia. *International Journal of Environmental Research and Public Health*. https://doi.org/10.3390/ijerph17218036

Ridder, D. D., Sandoval, J., Vuilleumier, N., Stringhini, S., Spechbach, H., Joost, S., Kaiser, L., & Guessous, I. (2020). Geospatial digital monitoring of COVID-19 cases at high spatiotemporal resolution. *The Lancet Digital Health*, *2*(8), e393–e394. https://doi.org/10.1016/S2589-7500(20)30139-4

Rincón Ruiz, A., Pascual, U., & Romero, M. (2013). An exploratory spatial analysis of illegal coca cultivation in Colombia using local indicators of spatial association and socioecological variables. *Ecological Indicators*. https://doi.org/10.1016/j.ecolind.2013.04.008

Rinner, C., & Hussain, M. (2011). Toronto's Urban Heat Island—Exploring the Relationship between Land Use and Surface Temperature. *Remote Sensing*, *3*(6), Article 6. https://doi.org/10.3390/rs3061251

Robertson, C., & Nelson, T. A. (2014). An Overview of Spatial Analysis of Emerging Infectious Diseases. *The Professional Geographer*, *66*(4), 579–588. https://doi.org/10.1080/00330124.2014.907702

Robertson, C., Nelson, T. A., MacNab, Y. C., & Lawson, A. B. (2010). Review of methods for space–time disease surveillance. *Spatial and Spatio-Temporal Epidemiology*, *1*(2), 105–116. https://doi.org/10.1016/j.sste.2009.12.001

Rohleder, S., & Bozorgmehr, K. (2021). Monitoring the spatiotemporal epidemiology of Covid-19 incidence and mortality: A small-area analysis in Germany. *Spatial and Spatio-Temporal Epidemiology*, *38*, 100433. https://doi.org/10.1016/j.sste.2021.100433

Rohr, J. R., & Cohen, J. M. (2020). Understanding how temperature shifts could impact infectious disease. *PLoS Biology*, *18*(11), e3000938. https://doi.org/10.1371/journal.pbio.3000938

Roosa, M. W., & White, R. M. B. (2014). Neighborhood Effects. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research* (pp. 4328–4331). Springer Netherlands. https://doi.org/10.1007/978-94-007-0753-5_1931

Rosillo, N., Del-Águila-Mejía, J., Rojas-Benedicto, A., Guerrero-Vadillo, M., Peñuelas, M., Mazagatos, C., Segú-Tell, J., Ramis, R., & Gómez-Barroso, D. (2021). Real time surveillance of COVID-19 space and time clusters during the summer 2020 in Spain. *BMC Public Health*, *21*(1), 961. https://doi.org/10.1186/s12889-021-10961-z

Runkle, J. D., Sugg, M. M., Leeper, R. D., Rao, Y., Matthews, J. L., & Rennie, J. J. (2020). Short-term effects of specific humidity and temperature on COVID-19 morbidity in select US cities. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.140093

Rydin, Y., Bleahu, A., Davies, M., Dávila, J. D., Friel, S., De Grandis, G., Groce, N., Hallal, P. C., Hamilton, I., Howden-Chapman, P., Lai, K.-M., Lim, C. J., Martins, J., Osrin, D., Ridley, I., Scott, I., Taylor, M., Wilkinson, P., & Wilson, J. (2012). Shaping cities for health: Complexity and the planning of urban environments in the 21st century. *Lancet*, *379*(9831), 2079–2108. https://doi.org/10.1016/S0140-6736(12)60435-8

Rytkönen, M. J. (2004). Not all maps are equal: GIS and spatial analysis in epidemiology. *International Journal of Circumpolar Health*. https://doi.org/10.3402/ijch.v63i1.17642

Ryu, E., Olson, J. E., Juhn, Y. J., Hathcock, M. A., Wi, C.-I., Cerhan, J. R., Yost, K. J., & Takahashi, P. Y. (2018). Association between an individual housing-based socioeconomic index and inconsistent self-reporting of health conditions: A prospective cohort study in the Mayo Clinic Biobank. *BMJ Open*, *8*(5), e020054. https://doi.org/10.1136/bmjopen-2017-020054

Saavedra, P., Santana, A., Bello, L., Pacheco, J.-M., & Sanjuán, E. (2021). A Bayesian spatio-temporal analysis of mortality rates in Spain: Application to the COVID-19 2020 outbreak. *Population Health Metrics*, *19*(1), 27. https://doi.org/10.1186/s12963-021-00259-y

Saffary, T., Adegboye, O. A., Gayawan, E., Elfaki, F., Kuddus, M. A., & Saffary, R. (2020). Analysis of COVID-19 Cases' Spatial Dependence in US Counties Reveals Health Inequalities. *Frontiers in Public Health*. https://doi.org/10.3389/fpubh.2020.579190

Sahar, L., Foster, S. L., Sherman, R. L., Henry, K. A., Goldberg, D. W., Stinchcomb, D. G., & Bauer, J. E. (2019). GIScience and cancer: State of the art and trends for cancer surveillance and epidemiology. *Cancer*, *125*(15), 2544–2560. https://doi.org/10.1002/cncr.32052

Sahu, K. K., & Kumar, R. (2020). Preventive and treatment strategies of COVID-19: From community to clinical trials. *Journal of Family Medicine and Primary Care*, *9*(5), 2149–2157. https://doi.org/10.4103/jfmpc.jfmpc_728_20

Sannigrahi, S., Pilla, F., Basu, B., & Basu, A. S. (2020). The overall mortality caused by COVID-19 in the European region is highly associated with demographic composition: A

spatial regression-based approach. *Sustainable Cities and Society*, *62*, 102418. https://doi.org/10.1016/j.scs.2020.102418

Sannigrahi, S., Pilla, F., Basu, B., Basu, A. S., & Molter, A. (2020a). Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustainable Cities and Society*. https://doi.org/10.1016/j.scs.2020.102418

Sannigrahi, S., Pilla, F., Basu, B., Basu, A. S., & Molter, A. (2020b). Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustainable Cities and Society*. https://doi.org/10.1016/j.scs.2020.102418

Sarkar, A., Liu, G., Jin, Y., Xie, Z., & Zheng, Z.-J. (2020). Public health preparedness and responses to the coronavirus disease 2019 (COVID-19) pandemic in South Asia: A situation and policy analysis. *Global Health Journal*, *4*(4), 121–132. https://doi.org/10.1016/j.glohj.2020.11.003

Scarpone, C., Brinkmann, S. T., Große, T., Sonnenwald, D., Fuchs, M., & Walker, B. B. (2020). A multimethod approach for county-scale geospatial analysis of emerging infectious diseases: A cross-sectional case study of COVID-19 incidence in Germany. *International Journal of Health Geographics*, *19*(1), 32. https://doi.org/10.1186/s12942-020-00225-1

Schærström, A. (2009). Disease Diffusion. *International Encyclopedia of Human Geography*, 222–233. https://doi.org/10.1016/B978-008044910-4.00330-8

Schmid, V. J., & Held, L. (2007). Bayesian Age-Period-Cohort Modeling and Prediction—BAMP. *Journal of Statistical Software*, *21*, 1–15. https://doi.org/10.18637/jss.v021.i08

Schneider, M. C., & Machado, G. (2018). Environmental and socioeconomic drivers in infectious disease. *The Lancet Planetary Health*, *2*(5), e198–e199. https://doi.org/10.1016/S2542-5196(18)30069-X

Scott, J., & Carrington, P. J. (2011). *The SAGE Handbook of Social Network Analysis*. SAGE Publications.

Sedgwick, P. (2015). Understanding the ecological fallacy. *BMJ*, *351*, h4773. https://doi.org/10.1136/bmj.h4773

Sekertekin, A., & Bonafoni, S. (2020). Land Surface Temperature Retrieval from Landsat 5, 7, and 8 over Rural Areas: Assessment of Different Retrieval Algorithms and Emissivity Models and Toolbox Implementation. *Remote Sensing*, *12*(2), Article 2. https://doi.org/10.3390/rs12020294

Senanayake, I. P., Welivitiya, W. D. D. P., & Nadeeka, P. M. (2013). Remote sensing based analysis of urban heat islands with vegetation cover in Colombo city, Sri Lanka using Landsat-7 ETM+ data. *Urban Climate*, *5*, 19–35. https://doi.org/10.1016/j.uclim.2013.07.004

Shadmi, E., Chen, Y., Dourado, I., Faran-Perach, I., Furler, J., Hangoma, P., Hanvoravongchai, P., Obando, C., Petrosyan, V., Rao, K. D., Ruano, A. L., Shi, L., de Souza, L. E., Spitzer-Shohat, S., Sturgiss, E., Suphanchaimat, R., Uribe, M. V., & Willems, S. (2020). Health equity and COVID-19: Global perspectives. *International Journal for Equity in Health*, *19*(1), 104. https://doi.org/10.1186/s12939-020-01218-z

Shariati, M., Jahangiri-rad, M., Mahmud Muhammad, F., & Shariati, J. (2020). Spatial Analysis of COVID-19 and Exploration of Its Environmental and Socio-Demographic Risk

Factors Using Spatial Statistical Methods: A Case Study of Iran TT -. *Hdqir*, *5*(3), 145–154. https://doi.org/10.32598/hdq.5.3.358.1

Shariati, M., Mesgari, T., Kasraee, M., & Jahangiri-rad, M. (2020). Spatiotemporal analysis and hotspots detection of COVID-19 using geographic information system (March and April, 2020). *Journal of Environmental Health Science and Engineering*. https://doi.org/10.1007/s40201-020-00565-x

Sharifi, A., & Khavarian-Garmsir, A. R. (2020). The COVID-19 pandemic: Impacts on cities and major lessons for urban planning, design, and management. *The Science of the Total Environment*, *749*, 142391. https://doi.org/10.1016/j.scitotenv.2020.142391

Shereen, M. A., Khan, S., Kazmi, A., Bashir, N., & Siddique, R. (2020). COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research*. https://doi.org/10.1016/j.jare.2020.03.005

Sherman, R. L., Henry, K. A., Tannenbaum, S. L., Feaster, D. J., Kobetz, E., & Lee, D. J. (2014). Applying Spatial Analysis Tools in Public Health: An Example Using SaTScan to Detect Geographic Targets for Colorectal Cancer Screening Interventions. *Preventing Chronic Disease*, *11*, E41. https://doi.org/10.5888/pcd11.130264

Shi, P., Dong, Y., Yan, H., Zhao, C., Li, X., Liu, W., He, M., Tang, S., & Xi, S. (2020). Impact of temperature on the dynamics of the COVID-19 outbreak in China. *Science of The Total Environment*, *728*, 138890. https://doi.org/10.1016/j.scitotenv.2020.138890

Shrestha, S., & Stopka, T. J. (2022). Spatial Epidemiology and Public Health. In F. S. Faruque (Ed.), *Geospatial Technology for Human Well-Being and Health* (pp. 49–77). Springer International Publishing. https://doi.org/10.1007/978-3-030-71377-5_4

Silva, A. P. de S. C., Holanda, E. R. de, Abreu, P. D. de, & Freitas, M. V. de A. (2021). COVID-19 in children in the state of Pernambuco: Spatial analysis of confirmed severe cases and the Human Development Index. *Revista Da Sociedade Brasileira De Medicina Tropical*, *54*, e0782-2020. https://doi.org/10.1590/0037-8682-0782-2020

Silverstein, W. K., Stroud, L., Cleghorn, G. E., & Leis, J. A. (2020). First imported case of 2019 novel coronavirus in Canada, presenting as mild pneumonia. *The Lancet*. https://doi.org/10.1016/S0140-6736(20)30370-6

Smereka, J., & Szarpak, L. (2020). COVID 19 a challenge for emergency medicine and every health care professional. *The American Journal of Emergency Medicine*, *38*(10), 2232–2233. https://doi.org/10.1016/j.ajem.2020.03.038

Smith, C. M., Le Comber, S. C., Fry, H., Bull, M., Leach, S., & Hayward, A. C. (2015). Spatial methods for infectious disease outbreak investigations: Systematic literature review. *Eurosurveillance*. https://doi.org/10.2807/1560-7917.ES.2015.20.39.30026

Snow, J. (1856). On the Mode of Communication of Cholera. *Edinburgh Medical Journal*, *1*(7), 668–670.

Snyder, B. F., & Parks, V. (2020). Spatial variation in socio-ecological vulnerability to Covid-19 in the contiguous United States. *Health and Place*. https://doi.org/10.1016/j.healthplace.2020.102471

Sobrino, J. A., Jiménez-Muñoz, J. C., & Paolini, L. (2004). Land surface temperature retrieval from LANDSAT TM 5. *Remote Sensing of Environment*. https://doi.org/10.1016/j.rse.2004.02.003

Sobrino, J. A., Jimenez-Munoz, J. C., Soria, G., Romaguera, M., Guanter, L., Moreno, J., Plaza, A., & Martinez, P. (2008). Land Surface Emissivity Retrieval From Different VNIR and TIR Sensors. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(2), 316–327. https://doi.org/10.1109/TGRS.2007.904834

Song, L., & Zhou, Y. (2020). The COVID-19 Pandemic and Its Impact on the Global Economy: What Does It Take to Turn Crisis into Opportunity? *China & World Economy*, *28*(4), 1–25. https://doi.org/10.1111/cwe.12349

Soucy, J.-P. R., Ghasemi, A., Sturrock, S. L., Berry, I., Buchan, S. A., MacFadden, D. R., Daneman, N., Gibb, N., & Brown, K. A. (2021). *Increased Interregional Travel to Shopping Malls and Restaurants in Response to Differential COVID-19 Restrictions in the Greater Toronto Area* (p. 2021.04.23.21255959). https://doi.org/10.1101/2021.04.23.21255959

Souza, E. C. de O., Santos, E. S. dos, Rosa, A. M., & Botelho, C. (2019). Space-time scan for identification of risk areas for hospitalization of children due to asthma in Mato Grosso, Brazil. *Revista Brasileira de Epidemiologia*, *22*. https://doi.org/10.1590/1980-549720190019

Statistics Canada. (2017). *Municipalities in Canada with the largest and fastest-growing populations between 2011 and 2016*. https://www12.statcan.gc.ca/census-recensement/2016/as-sa/98-200-x/2016001/98-200-x2016001-eng.cfm

Statistics Canada. (2021). *Census tract (CT)—Census Dictionary*. https://www12.statcan.gc.ca/census-recensement/2011/ref/dict/geo013-eng.cfm

Steele, L., Orefuwa, E., Bino, S., Singer, S. R., Lutwama, J., & Dickmann, P. (2020). Earlier Outbreak Detection—A Generic Model and Novel Methodology to Guide Earlier Detection Supported by Data From Low- and Mid-Income Countries. *Frontiers in Public Health*, *8*. https://www.frontiersin.org/articles/10.3389/fpubh.2020.00452

Steinbrook, R. (2020). Contact Tracing, Testing, and Control of COVID-19—Learning From Taiwan. *JAMA Internal Medicine*, *180*(9), 1163–1164. https://doi.org/10.1001/jamainternmed.2020.2072

Stringhini, S., Carmeli, C., Jokela, M., Avendaño, M., McCrory, C., d'Errico, A., Bochud, M., Barros, H., Costa, G., Chadeau-Hyam, M., Delpierre, C., Gandini, M., Fraga, S., Goldberg, M., Giles, G. G., Lassale, C., Kenny, R. A., Kelly-Irving, M., Paccaud, F., … Kivimäki, M. (2018). Socioeconomic status, non-communicable disease risk factors, and walking speed in older adults: Multi-cohort population based study. *BMJ*, *360*, k1046. https://doi.org/10.1136/bmj.k1046

Stroup, D. F., Wharton, M., Kafadar, K., & Dean, A. G. (1993). Evaluation of a Method for Detecting Aberrations in Public Health Surveillance Data. *American Journal of Epidemiology*, *137*(3), 373–380. https://doi.org/10.1093/oxfordjournals.aje.a116684

Strully, K., Yang, T.-C., & Liu, H. (2021). Regional variation in COVID-19 disparities: Connections with immigrant and Latinx communities in U.S. counties. *Annals of Epidemiology*, *53*, 56-62.e2. https://doi.org/10.1016/j.annepidem.2020.08.016

Sun, F., Matthews, S. A., Yang, T. C., & Hu, M. H. (2020a). A spatial analysis of the COVID-19 period prevalence in U.S. counties through June 28, 2020: Where geography matters? *Annals of Epidemiology*. https://doi.org/10.1016/j.annepidem.2020.07.014

Sun, F., Matthews, S. A., Yang, T. C., & Hu, M. H. (2020b). A spatial analysis of the COVID-19 period prevalence in U.S. counties through June 28, 2020: Where geography matters? *Annals of Epidemiology*. https://doi.org/10.1016/j.annepidem.2020.07.014

Sun, Y., Hu, X., & Xie, J. (2020a). Spatial inequalities of COVID-19 mortality rate in relation to socioeconomic and environmental factors across England. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.143595

Sun, Y., Hu, X., & Xie, J. (2020b). Spatial inequalities of COVID-19 mortality rate in relation to socioeconomic and environmental factors across England. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.143595

Suryowati, K., Bekti, R. D., & Faradila, A. (2018). A Comparison of Weights Matrices on Computation of Dengue Spatial Autocorrelation. *IOP Conference Series: Materials Science and Engineering*, *335*, 012052. https://doi.org/10.1088/1757-899X/335/1/012052

Suthar, A. B., Wang, J., Seffren, V., Wiegand, R. E., Griffing, S., & Zell, E. (2022). Public health impact of covid-19 vaccines in the US: Observational study. *BMJ*, *377*, e069317. https://doi.org/10.1136/bmj-2021-069317

Sy, K. T. L., White, L. F., & Nichols, B. E. (2021). Population density and basic reproductive number of COVID-19 across United States counties. *PLOS ONE*, *16*(4), e0249271. https://doi.org/10.1371/journal.pone.0249271

Tabish, S. A. (2020). Covid-19 Pandemic: Emerging Perspectives and Future Trends. *Journal of Public Health Research*, *9*(1), jphr.2020.1786. https://doi.org/10.4081/jphr.2020.1786

Tam, B. Y., Gough, W. A., & Mohsin, T. (2015). The impact of urbanization and the urban heat island effect on day to day temperature variation. *Urban Climate*, *12*, 1–10. https://doi.org/10.1016/j.uclim.2014.12.004

Thayn, J. B., & Simanis, J. M. (2013). Accounting for Spatial Autocorrelation in Linear Regression Models Using Spatial Filtering with Eigenvectors. *Annals of the Association of American Geographers*. https://doi.org/10.1080/00045608.2012.685048

Threats, I. of M. (US) F. on M., Knobler, S., Mahmoud, A., Lemon, S., & Pray, L. (2006). Summary and Assessment. In *The Impact of Globalization on Infectious Disease Emergence and Control: Exploring the Consequences and Opportunities: Workshop Summary*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK56579/

Tiefelsdorf, M., & Kyunghee Rhyu. (2022). *Disparities in COVID-19 Vaccination Rates among the Counties of Texas*. http://www.spatialfiltering.com/ThinkR/Downloads/VacModelTX.html#23_Rates

Tiwari, P. K., Rai, R. K., Khajanchi, S., Gupta, R. K., & Misra, A. K. (2021). Dynamics of coronavirus pandemic: Effects of community awareness and global information campaigns. *The European Physical Journal Plus*, *136*(10), Article 10. https://doi.org/10.1140/epjp/s13360-021-01997-6

Tobías, A., & Molina, T. (2020). Is temperature reducing the transmission of COVID-19 ? *Environmental Research*. https://doi.org/10.1016/j.envres.2020.109553

Toronto lockdown—One of the world's longest? (2021, May 23). *BBC News*. https://www.bbc.com/news/world-us-canada-57079577

Travaglio, M., Yu, Y., Popovic, R., Selley, L., Leal, N. S., & Martins, L. M. (2021). Links between air pollution and COVID-19 in England. *Environmental Pollution*, *268*, 115859. https://doi.org/10.1016/j.envpol.2020.115859

Tsutakawa, R. K., Shoop, G. L., & Marienfeld, C. J. (1985). Empirical bayes estimation of cancer mortality rates. *Statistics in Medicine*, *4*(2), 201–212. https://doi.org/10.1002/sim.4780040210

Tyrovolas, S., Giné-Vázquez, I., Fernández, D., Morena, M., Koyanagi, A., Janko, M., Haro, J. M., Lin, Y., Lee, P., Pan, W., Panagiotakos, D., & Molassiotis, A. (2021). Estimating the COVID-19 Spread Through Real-time Population Mobility Patterns: Surveillance in Low- and Middle-Income Countries. *Journal of Medical Internet Research*, *23*(6), e22999. https://doi.org/10.2196/22999

Ullah, S., Nor, N. H. M., Daud, H., Zainuddin, N., Gandapur, M. S. J., Ali, I., & Khalil, A. (2021). Spatial cluster analysis of COVID-19 in Malaysia (Mar-Sep, 2020). *Geospatial Health*, *16*(1), Article 1. https://doi.org/10.4081/gh.2021.961

Unkel, S., Farrington, C. P., Garthwaite, P. H., Robertson, C., & Andrews, N. (2012). Statistical methods for the prospective detection of infectious disease outbreaks: A review. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *175*(1), 49–82. https://doi.org/10.1111/j.1467-985X.2011.00714.x

USGS- Landsat. (2021). https://www.usgs.gov/core-science-systems/nli/landsat

van Seventer, J. M., & Hochberg, N. S. (2017). Principles of Infectious Diseases: Transmission, Diagnosis, Prevention, and Control. *International Encyclopedia of Public Health*, 22–39. https://doi.org/10.1016/B978-0-12-803678-5.00516-6

Vaz, E. (2021). COVID-19 in Toronto: A Spatial Exploratory Analysis. *Sustainability*, *13*(2), Article 2. https://doi.org/10.3390/su13020498

Vehtari, A., & Ojanen, J. (2012). A survey of Bayesian predictive methods for model assessment, selection and comparison. *Statistics Surveys*, *6*(none), 142–228. https://doi.org/10.1214/12-SS102

Vehtari, A., Simpson, D. P., Yao, Y., & Gelman, A. (2019). Limitations of "Limitations of Bayesian Leave-one-out Cross-Validation for Model Selection". *Computational Brain & Behavior*, *2*(1), 22–27. https://doi.org/10.1007/s42113-018-0020-6

Verma, R., Yabe, T., & Ukkusuri, S. V. (2021). Spatiotemporal contact density explains the disparity of COVID-19 spread in urban neighborhoods. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-90483-1

Wachtler, B., Michalski, N., Nowossadeck, E., Diercke, M., Wahrendorf, M., Santos-Hövener, C., Lampert, T., & Hoebel, J. (2020). *Socioeconomic inequalities and COVID-19 – A review of the current international literature*. https://doi.org/10.25646/7059

Wade, L. (2020). An unequal blow. *Science (New York, N.Y.)*, *368*(6492), 700–703. https://doi.org/10.1126/science.368.6492.700

Wagner, M. M., Tsui, F. C., Espino, J. U., Dato, V. M., Sittig, D. F., Caruana, R. A., McGinnis, L. F., Deerfield, D. W., Druzdzel, M. J., & Fridsma, D. B. (2001). The emerging science of very early detection of disease outbreaks. *Journal of Public Health Management and Practice: JPHMP*, *7*(6), 51–59. https://doi.org/10.1097/00124784-200107060-00006

Wakefield, J. (2003). Sensitivity Analyses for Ecological Regression. *Biometrics*, *59*(1), 9–17. https://doi.org/10.1111/1541-0420.00002

Wang, L., Xu, C., Wang, J., Qiao, J., Yan, M., & Zhu, Q. (2021). Spatiotemporal heterogeneity and its determinants of COVID-19 transmission in typical labor export provinces of China. *BMC Infectious Diseases, 21*(1), 242. https://doi.org/10.1186/s12879-021-05926-x

Wang, P., Hu, T., Liu, H., & Zhu, X. (2022). Exploring the impact of under-reported cases on the COVID-19 spatiotemporal distributions using healthcare workers infection data. *Cities (London, England), 123*, 103593. https://doi.org/10.1016/j.cities.2022.103593

Wang, P., Ren, H., Zhu, X., Fu, X., Liu, H., & Hu, T. (2021). Spatiotemporal characteristics and factor analysis of SARS-CoV-2 infections among healthcare workers in Wuhan, China. *The Journal of Hospital Infection*, *110*, 172–177. https://doi.org/10.1016/j.jhin.2021.02.002

Wang, S., Ma, Q., Ding, H., & Liang, H. (2018). Detection of urban expansion and land surface temperature change using multi-temporal landsat images. *Resources, Conservation and Recycling*, *128*, 526–534. https://doi.org/10.1016/j.resconrec.2016.05.011

Wang, X., Wang, L., Zhang, X., & Fan, F. (2022). The spatiotemporal evolution of COVID-19 in China and its impact on urban economic resilience. *China Economic Review*, *74*, 101806. https://doi.org/10.1016/j.chieco.2022.101806

Wang, Y., Berardi, U., & Akbari, H. (2016). Comparing the effects of urban heat island mitigation strategies for Toronto, Canada. *Energy and Buildings*, *114*, 2–19. https://doi.org/10.1016/j.enbuild.2015.06.046

Wang, Y., Liu, Y., Struthers, J., & Lian, M. (2020). Spatiotemporal Characteristics of the COVID-19 Epidemic in the United States. *Clinical Infectious Diseases*. https://doi.org/10.1093/cid/ciaa934

Watson, O. J., Barnsley, G., Toor, J., Hogan, A. B., Winskill, P., & Ghani, A. C. (2022). Global impact of the first year of COVID-19 vaccination: A mathematical modelling study. *The Lancet Infectious Diseases*, *22*(9), 1293–1302. https://doi.org/10.1016/S1473-3099(22)00320-6

Weng, Q., Fu, P., & Gao, F. (2014). Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sensing of Environment*, *145*, 55–67. https://doi.org/10.1016/j.rse.2014.02.003

Wheeler, D. C. (2007). Diagnostic tools and a remedial method for collinearity in geographically weighted regression. *Environment and Planning A*. https://doi.org/10.1068/a38325

Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*. https://doi.org/10.1007/s10109-005-0155-6

WHO. (2021). *World Health Organization:COVID-19 dashboard*. https://covid19.who.int/

World Health Organization. (2020). *WHO COVID-19 Case definition*. https://www.who.int/publications/i/item/WHO-2019-nCoV-Surveillance_Case_Definition-2020.2

Wu, X., Yin, J., Li, C., Xiang, H., Lv, M., & Guo, Z. (2020). Natural and human environment interactively drive spread pattern of COVID-19: A city-level modeling study in China. *Science of The Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.143343

Wu, X., & Zhang, J. (2021). Exploration of spatial-temporal varying impacts on COVID-19 cumulative case in Texas using geographically weighted regression (GWR).

Environmental Science and Pollution Research International, *28*(32), 43732–43746. https://doi.org/10.1007/s11356-021-13653-8

Wu, Y., Yan, X., Zhao, S., Wang, J., Ran, J., Dong, D., Wang, M., Fung, H., Yeoh, E. kiong, & Chung, R. Y. N. (2020). Association of time to diagnosis with socioeconomic position and geographical accessibility to healthcare among symptomatic COVID-19 patients: A retrospective study in Hong Kong. *Health and Place*. https://doi.org/10.1016/j.healthplace.2020.102465

Xie, Z., Qin, Y., Li, Y., Shen, W., Zheng, Z., & Liu, S. (2020). Spatial and temporal differentiation of COVID-19 epidemic spread in mainland China and its influencing factors. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.140929

Xie, Z., Zhao, R., Ding, M., & Zhang, Z. (2021). A Review of Influencing Factors on Spatial Spread of COVID-19 Based on Geographical Perspective. *International Journal of Environmental Research and Public Health*, *18*(22), Article 22. https://doi.org/10.3390/ijerph182212182

Xiong, Y., Wang, Y., Chen, F., & Zhu, M. (2020). Spatial statistics and influencing factors of the COVID-19 epidemic at both prefecture and county levels in Hubei Province, China. *International Journal of Environmental Research and Public Health*. https://doi.org/10.3390/ijerph17113903

Xu, F., & Beard, K. (2021). A comparison of prospective space-time scan statistics and spatiotemporal event sequence based clustering for COVID-19 surveillance. *PLOS ONE*, *16*(6), e0252990. https://doi.org/10.1371/journal.pone.0252990

Yandell, B. S., & Anselin, L. (1990). Spatial Econometrics: Methods and Models. *Journal of the American Statistical Association*. https://doi.org/10.2307/2290042

Yang, T.-C., Kim, S., Zhao, Y., & Choi, S.-W. E. (2021). Examining spatial inequality in COVID-19 positivity rates across New York City ZIP codes. *Health & Place*, *69*, 102574. https://doi.org/10.1016/j.healthplace.2021.102574

Ye, L., & Hu, L. (2020). Spatiotemporal distribution and trend of COVID-19 in the Yangtze river Delta region of the People's Republic of China. *Geospatial Health*. https://doi.org/10.4081/gh.2020.889

Ye, Y., & Qiu, H. (2021). Using urban landscape pattern to understand and evaluate infectious disease risk. *Urban Forestry & Urban Greening*, *62*, 127126. https://doi.org/10.1016/j.ufug.2021.127126

Yin, C. L., Meng, F., & Yu, Q. R. (2020). Calculation of land surface emissivity and retrieval of land surface temperature based on a spectral mixing model. *Infrared Physics & Technology*, *108*, 103333. https://doi.org/10.1016/j.infrared.2020.103333

Yoshikawa, Y., & Kawachi, I. (2021). Association of Socioeconomic Characteristics With Disparities in COVID-19 Outcomes in Japan. *JAMA Network Open*, *4*(7), e2117060. https://doi.org/10.1001/jamanetworkopen.2021.17060

You, H., Wu, X., & Guo, X. (2020). Distribution of covid-19 morbidity rate in association with social and economic factors in wuhan, china: Implications for urban development. *International Journal of Environmental Research and Public Health*. https://doi.org/10.3390/ijerph17103417

Yu, D., Peterson, N. A., Sheffer, M. A., Reid, R. J., & Schnieder, J. E. (2010). Tobacco outlet density and demographics: Analysing the relationships with a spatial regression approach. *Public Health*. https://doi.org/10.1016/j.puhe.2010.03.024

Yu, H., Fotheringham, A. S., Li, Z., Oshan, T., Kang, W., & Wolf, L. J. (2020). Inference in Multiscale Geographically Weighted Regression. *Geographical Analysis*, *52*(1), 87–106. https://doi.org/10.1111/gean.12189

Yu, H., Li, J., Bardin, S., Gu, H., & Fan, C. (2021). Spatiotemporal Dynamic of COVID-19 Diffusion in China: A Dynamic Spatial Autoregressive Model Analysis. *ISPRS International Journal of Geo-Information*, *10*(8), Article 8. https://doi.org/10.3390/ijgi10080510

Yu, X., Guo, X., & Wu, Z. (2014). Land Surface Temperature Retrieval from Landsat 8 TIRS—Comparison between Radiative Transfer Equation-Based Method, Split Window Algorithm and Single Channel Method. *Remote Sensing*, *6*(10), Article 10. https://doi.org/10.3390/rs6109829

Zajacova, A., Jehn, A., Stackhouse, M., Denice, P., & Ramos, H. (2020). Changes in health behaviours during early COVID-19 and socio-demographic disparities: A cross-sectional analysis. *Canadian Journal of Public Health*, *111*(6), 953–962. https://doi.org/10.17269/s41997-020-00434-y

Zhang, J., Wu, X., & Chow, T. E. (2021). Space-Time Cluster's Detection and Geographical Weighted Regression Analysis of COVID-19 Mortality on Texas Counties. *International Journal of Environmental Research and Public Health*, *18*(11), 5541. https://doi.org/10.3390/ijerph18115541

Zhang, X., Rao, H., Wu, Y., Huang, Y., & Dai, H. (2020). Comparison of spatiotemporal characteristics of the COVID-19 and SARS outbreaks in mainland China. *BMC Infectious Diseases, 20*(1), 805. https://doi.org/10.1186/s12879-020-05537-y

Zhang, Y., & Sun, L. (2019). Spatial-temporal impacts of urban land use land cover on land surface temperature: Case studies of two Canadian urban areas. *International Journal of Applied Earth Observation and Geoinformation, 75*, 171–181. https://doi.org/10.1016/j.jag.2018.10.005

Zhao, S., Qin, Q., Yang, Y., Xiong, Y., & Qiu, G. (2009). Comparison of two split-window methods for retrieving land surface temperature from MODIS data. *Journal of Earth System Science, 118*(4), 345. https://doi.org/10.1007/s12040-009-0027-4

Zheng, J., Shen, G., Hu, S., Han, X., Zhu, S., Liu, J., He, R., Zhang, N., Hsieh, C.-W., Xue, H., Zhang, B., Shen, Y., Mao, Y., & Zhu, B. (2022). Small-scale spatiotemporal epidemiology of notifiable infectious diseases in China: A systematic review. *BMC Infectious Diseases, 22*(1), 723. https://doi.org/10.1186/s12879-022-07669-9

Zhou, Y., Wang, L., Zhang, L., Shi, L., Yang, K., He, J., Zhao, B., Overton, W., Purkayastha, S., & Song, P. (2020). A Spatiotemporal Epidemiological Prediction Model to Inform County-Level COVID-19 Risk in the United States. *Harvard Data Science Review*. https://doi.org/10.1162/99608f92.79e1f45e

Appendices

Appendix A: (Chapter 4)

Appendix 1: Land Surface Temperature (LST) Retrieval Method

Landsat Mission

The Landsat program, hosted by the USGS Earth Resources Observation and Science (EROS) centre, is the longest-running enterprise for the acquisition of satellite imagery of Earth (*USGS- Landsat*, 2021). For our study, we have used Landsat Collection 2 level 1 dataset. Landsat 8, the most recently launched Landsat satellite, has two sensors: the Operational Land Imager (OLI) and the Thermal Infrared (TIR). Landsat 8 and Landsat 7 satellites have a 16-day repeat cycle and circle the earth in a sun-synchronous, near-polar orbit at an altitude of 705 km, inclined at 98.2 degrees. The satellites circle the earth's orbit every 99 minutes with an equatorial crossing time: 10:00 AM ±15 minutes. Landsat 8 images consist of nine spectral bands with a spatial resolution of 30 meters for Bands 1 to 7 and 9, and 10 and 11 (thermal) are collected at 100 meters (*USGS- Landsat*, 2021). Landsat 7 images consist of eight spectral bands with a spatial resolution of 30 meters for Bands 1 to 7, while Band 6 (thermal) collects both high and low gain (60/30 meters) for all scenes.

Land Surface Temperature (LST) Retrieval Method

In the last few decades, several methodologies with improved algorithms for computing land surface temperature using satellite-based thermal infrared (TIR) data has been developed (Coll & Caselles, 1997; Jiménez-Muñoz et al., 2014; Weng et al., 2014). A few studies comparing these methods found that Sobrino et al. 2008's Land Surface Emissivity (LSE) model(Sobrino et al., 2008) provided the highest accuracy for extracting land surface temperature using the Landsat imagery (Neinavaz et al., 2020; Sekertekin & Bonafoni, 2020; Yin et al., 2020). We applied an automated extraction toolbox in ArcGIS Desktop developed by Aliiahsan et al. (Sekertekin & Bonafoni, 2020); that computes Land Surface Temperature (LST) using the Radiative Transfer Equation (RTE) method based on Sobrino's LSE model (Sobrino et al., 2008).

NDVI Threshold ($NDVI^{THM}$)-Based LSE Model

We used the NDVI Thresholds method proposed by Sobrino et al. in 2008 (Sobrino et al., 2008) to compute the Land Surface Emissivity (LSE) from NDVI threshold values considering three different cases:1) NDVI <0.2, the pixel is considered as bare soil, and the emissivity is obtained from the reflectance values in the red region, 2) $0.2 \leq NDVI \geq 0.5$, the pixel is composed of a mixture of bare soil and vegetation, 3) NDVI > 0.5, the pixels with NDVI values higher than 0.5 are considered as fully vegetated areas(Sobrino et al., 2008). The model is expressed in the following equation:

$$\varepsilon = \begin{cases} a_i \rho_R + b_i & NDVI < 0.2 \\ \varepsilon_v + \varepsilon_v(1 - P_v) + d\varepsilon, d\varepsilon = (1 - \varepsilon_s)(1 - P_v) F \varepsilon_v & 0.2 \leq NDVI \geq 0.5 \\ \varepsilon_v + d\varepsilon & NDVI > 0.5 \end{cases}$$

ε_v is the soil and ε_s is the vegetation emissivity. $d\varepsilon$ is the cavity effect due to surface roughness, F is a geometrical shape factor assumed as the mean value of 0.55 (Sobrino et al., 2004), ρ_R is the reflectance value of the red band, a_ii and b_i are estimated from an empirical relationship between the red band reflectance and Moderate Resolution Imaging Spectroradiometer (MODIS) emissivity library.

LST calculation using radiative transfer equation (RTE)

For the RTE-based method, we have used an atmospheric correction parameter calculator (NASA, 2021) to extract global atmospheric profiles for each image date, time and location. The correction parameter used the NCEP (National Centre for Environmental Prediction) to simulate atmospheric transmittance, upwelling and downwelling radiances[7]. Removing the effects of the atmosphere in the thermal region is the essential step necessary to use the thermal band as atmospheric correction allows avoiding systematic errors in the predicted surface temperature[7]. With these parameters, the space-reaching radiance was converted to a surface-leaving radiance[7].

$$L_{TOA} = \tau\varepsilon L_T + L_u + (1-\varepsilon)L_d$$

Where τ is the atmospheric transmission, ε is the emissivity of the surface, L_T is the radiance of a blackbody target of kinetic temperature T, L_u is the upwelling or atmospheric path radiance, L_d is the downwelling or sky radiance, and L_{TOA} is TOA radiance measured by the instrument.

LST was retrieved using a single TIR band with the inversion of the radiative transfer equation (RTE) according to the following equation where L_λ^{sen} ($W{\cdot}m^{-2}{\cdot}sr^{-1}{\cdot}\mu m^{-1}$) is at-sensor registered radiance of the related thermal band, B_λ ($W{\cdot}m^{-2}{\cdot}sr^{-1}{\cdot}\mu m^{-1}$) is the blackbody radiance.

$$L_\lambda^{sen} = [\varepsilon B_\lambda(T_S) + (1-\varepsilon)L_\lambda^{\downarrow}]\tau + L_\lambda^{\uparrow}$$

Blackbody radiance (B_λ) at a temperature of can be obtained by inverting the previous equation:

$$B_\lambda(T_S) = \frac{L_\lambda^{sen} - L_\lambda^{\uparrow} - \tau(1-\varepsilon)L_\lambda^{\downarrow}}{\tau\varepsilon}$$

Land Surface Temperature T_S can be obtained by inverting Planck's Law as:

$$T_S = \frac{K_2}{\ln\left[\frac{K_1}{\frac{L_\lambda^{sen} - L_\lambda^{\uparrow} - \tau(1-\varepsilon)L_\lambda^{\downarrow}}{\tau\varepsilon}} + 1\right]}$$

where K_1 and K_2 are calibration constants for Landsat data reported.

Retrieval of spectral radiance and Brightness temperature (*T*) retrieval

The Top of Atmosphere (TOA) spectral radiance value is calculated for Landsat 7 using the following equation:

$$L_\lambda = \left[\frac{LMAX_\lambda - LMIN_\lambda}{QCALMAX - QCALMIN}\right] x\,[Q_{CAL} - QCALMIN] + LMIN_\lambda$$

where L_λ is TOA spectral radiance (Watts/(m^2·srad·μm)), Q_{CAL} is the quantized calibrated pixel value in DN, $LMIN_\lambda$ (Watts/(m^2·srad·μm)) is the spectral radiance scaled to $QCALMIN$, $LMAX_\lambda$ (Watts/(m^2·srad·μm)) is the spectral radiance scaled to $QCALMAX$, $QCALMIN$ is the minimum quantized calibrated pixel value in DN and $QCALMIN$ is the maximum quantized calibrated pixel value in DN. $LMIN_\lambda$, $LMAX_\lambda$, $QCALMIN$, and $QCALMAX$ values are obtained from the metadata file of Landsat 7 data(Sekertekin & Bonafoni, 2020).

The TOA spectral radiance value is calculated for Landsat 8 using the following equation.

$$L_\lambda = M_L\,.\,Q_{CAL} + A_L$$

where L_λ is the TOA spectral radiance (Watts/(m^2·srad·μm)), M_L is the band-specific multiplicative rescaling factor from the metadata, A_L is the band-specific additive rescaling factor from the metadata, Q_{CAL} is the quantized and calibrated standard product pixel values (DN).

The brightness temperature T can be generated by the following equation:

$$T = \frac{K_2}{\ln\left(\frac{K_1}{L} + 1\right)}$$

where T refers to the effective at-satellite brightness temperature in Kelvin, K_1 (Watts/(m^2·srad·μm)) and K_2 (Kelvin) are the calibration constants and $L_\lambda$is the spectral radiance.

Table S4.1. Parameters for the Land Surface Temperature (LST) calculation.

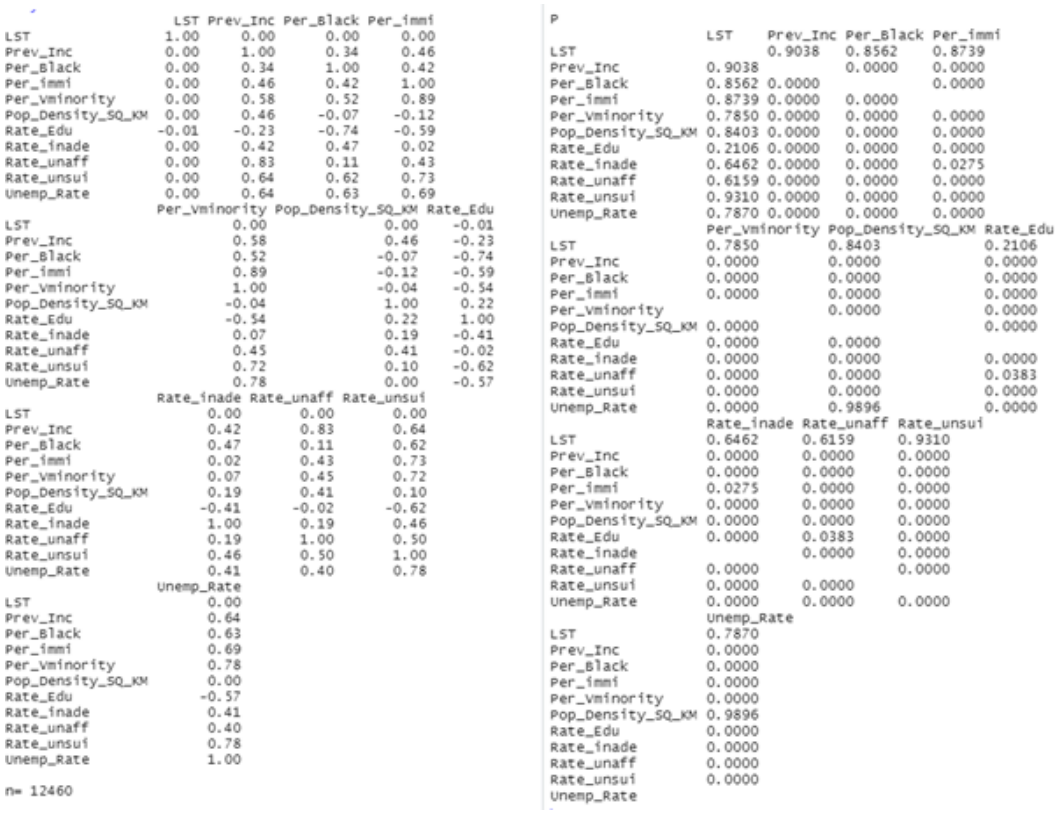
			Atmospheric Correction Parameter Calculator (https://atmcorr.gsfc.nasa.gov/cgi-bin/atm_corr.pl)						
Week	**Date of image**	**Scene Center Time (GMT)**	**Landsat**	**Sun Elevation (From Metadata)**	**Atmospheric Transmission**	**Upwelling radiance**	**Downwelling radiance**	**Solar Zenith Angle (For Landsat 7)**	**Earth-Sun Distance (From metadata)**
1	24-Jan-20	16:03:55	Landsat 8	24.35807661	0.91	0.54	0.91		
2	01-Feb-20	15:43:06	Landsat 7	24.49170863	0.92	0.38	0.64	65.5082914	0.9853188
3	02-Feb-20	15:57:41	Landsat 8	26.50548037	0.92	0.43	0.74		
4	09-Feb-20	16:03:50	Landsat 8	28.49063776	0.94	0.29	0.5		
5	18-Feb-20	15:57:38	Landsat 8	31.37765848	0.91	0.52	0.88		
6	25-Feb-20	16:03:47	Landsat 8	33.8380565	0.89	0.59	1		
7	05-Mar-20	15:57:32	Landsat 8	37.19903778	0.96	0.21	0.37		
8	12-Mar-20	16:03:40	Landsat 8	39.91607	0.91	0.55	0.93		
9	21-Mar-20	15:57:25	Landsat 8	43.46096552	0.98	0.1	0.17		
10	28-Mar-20	16:03:31	Landsat 8	46.2074895	0.79	1.36	2.2		
11	29-Mar-20	15:33:52	Landsat 7	44.12493545	0.69	2.15	3.35	45.8750646	0.9985865
12	06-Apr-20	15:57:16	Landsat 8	49.64471163	0.94	0.34	0.59		
13	13-Apr-20	16:03:24	Landsat 8	52.19555246	0.74	1.73	2.79		
14	22-Apr-20	15:57:09	Landsat 8	55.24010206	0.97	0.15	0.26		
15	29-Apr-20	16:03:16	Landsat 8	57.37985022	0.76	1.62	2.63		
16	08-May-20	15:57:00	Landsat 8	59.7895888	0.96	0.22	0.38		
17	15-May-20	16:03:12	Landsat 8	61.35021213	0.63	2.55	3.97		
18	23-May-20	15:36:46	Landsat 7	59.21636685	0.7	2.15	3.4	30.7836332	1.0126091
19	24-May-20	15:57:03	Landsat 8	62.90624602	0.63	2.8	4.36		
20	31-May-20	16:03:17	Landsat 8	63.75394942	0.92	0.52	0.9		
21	09-Jun-20	15:57:12	Landsat 8	64.37367127	0.71	2.37	3.79		
22	16-Jun-20	16:03:27	Landsat 8	64.50220934	0.86	1.05	1.76		
23	25-Jun-20	15:57:21	Landsat 8	64.24295184	0.78	1.58	2.62		
24	02-Jul-20	16:03:35	Landsat 8	63.74261212	0.65	2.83	4.48		
25	11-Jul-20	15:57:28	Landsat 8	62.75393472	0.51	3.78	5.7		
26	18-Jul-20	16:03:41	Landsat 8	61.73733148	0.64	2.99	4.68		
27					*Data Unavailable*				
28	27-Jul-20	15:57:32	Landsat 8	60.14323089	0.47	4.33	6.37		
29	03-Aug-20	16:03:45	Landsat 8	58.6961457	0.63	2.82	4.4		
30	12-Aug-20	15:57:36	Landsat 8	56.58795429	0.83	1.36	2.27		
31	19-Aug-20	16:03:50	Landsat 8	54.76228155	0.83	1.22	2.02		

32	28-Aug-20	15:57:44	Landsat 8	52.2044129	0.7	2.32	3.68		
33	04-Sep-20	16:03:58	Landsat 8	50.06588993	0.83	1.2	2		
34	12-Sep-20	15:29:52	Landsat 7	44.03172077	0.72	1.9	3.04	45.9682792	1.0062831
35	13-Sep-20	15:57:51	Landsat 8	47.1568631	0.69	2.42	3.82		
36	20-Sep-20	16:04:04	Landsat 8	44.79112835	0.9	0.67	1.13		
37	29-Sep-20	15:57:55	Landsat 8	41.67289377	0.84	1.13	1.89		
38	06-Oct-20	16:04:07	Landsat 8	39.21883602	0.79	1.4	2.3		
39	15-Oct-20	15:57:57	Landsat 8	36.08716502	0.78	1.43	2.35		
40	22-Oct-20	16:04:07	Landsat 8	33.70881838	0.71	1.98	3.13		
41	31-Oct-20	15:57:56	Landsat 8	30.80408106	0.95	0.27	0.47		
42	*Data Unavailable*								
43	08-Nov-20	15:19:57	Landsat 7	25.8407203	0.89	0.83	1.37	64.1592797	0.9906193
44	16-Nov-20	15:57:54	Landsat 8	26.28530259	0.94	0.34	0.59		
45	23-Nov-20	16:04:07	Landsat 8	24.66576502	0.92	0.44	0.76		
46	02-Dec-20	15:57:58	Landsat 8	22.99062118	0.96	0.21	0.36		
47	09-Dec-20	16:04:09	Landsat 8	22.03707882	0.9	0.54	0.92		
48	18-Dec-20	15:57:57	Landsat 8	21.30569479	0.96	0.19	0.33		
49	25-Dec-20	16:04:06	Landsat 8	21.14419286	0.94	0.31	0.52		
50	02-Jan-21	15:22:10	Landsat 7	18.13418977	0.92	0.48	0.79	71.8658102	0.9833016
51	03-Jan-21	15:57:52	Landsat 8	21.47295072	0.89	0.6	1.02		
52	10-Jan-21	16:04:00	Landsat 8	22.14366309	0.95	0.27	0.47		
53	19-Jan-21	15:57:45	Landsat 8	22.14366309	0.95	0.22	0.39		
54	26-Jan-21	16:03:56	Landsat 8	24.96601689	0.91	0.46	0.78		
55	04-Feb-21	15:57:44	Landsat 8	27.25849939	0.96	0.18	0.32		
56	11-Feb-21	16:03:53	Landsat 8	29.34081589	0.97	0.12	0.21		
57	20-Feb-21	15:57:39	Landsat 8	32.3300296	0.97	0.16	0.28		
58	*Data Unavailable*								
59	28-Feb-21	15:11:48	Landsat 7	30.49313654	0.87	0.75	1.25	59.5068635	0.9906971
60	08-Mar-21	16:04:09	Landsat 8	22.03707882	0.91	0.52	0.88		
61	15-Mar-21	16:03:38	Landsat 8	41.00473898	0.97	0.14	0.24		
62	24-Mar-21	15:57:25	Landsat 8	44.55284896	0.88	0.7	1.18	45.447151	1.0058369
63	31-Mar-21	16:03:34	Landsat 8	47.28224338	0.87	0.76	1.27		
64	09-Apr-21	15:57:20	Landsat 8	50.67232025	0.74	1.8	2.89		
65	16-Apr-21	16:03:28	Landsat 8	53.16767185	0.87	0.74	1.24		
66	24-Apr-21	15:13:30	Landsat 7	49.52985399	0.88	0.7	1.18	40.470146	1.0058369
67	25-Apr-21	15:57:13	Landsat 8	56.11881991	0.92	0.49	0.83		
68	02-May-21	16:03:19	Landsat 8	58.16904696	0.74	1.75	2.84		
69	11-May-21	15:57:09	Landsat 8	60.45369328	0.92	0.44	0.76		
70	18-May-21	16:03:25	Landsat 8	61.89629835	0.81	1.45	2.41		

71	27-May-21	15:57:19	Landsat 8	63.29413436	0.9	0.62	1.07		
72	03-Jun-21	16:03:33	Landsat 8	64.01577194	0.67	2.44	3.83		
73	12-Jun-21	15:57:26	Landsat 8	64.47412564	0.72	2.14	3.42		
74	19-Jun-21	16:03:39	Landsat 8	64.48534021	0.76	1.88	3.09		
75	20-Jun-21	15:02:34	Landsat 7	56.08455683	0.69	2.26	3.55	33.9154432	1.0162019
76	28-Jun-21	15:57:30	Landsat 8	64.09149511	0.52	3.79	5.72		
77	05-Jul-21	16:03:42	Landsat 8	63.49402232	0.62	3.07	4.77		
78	14-Jul-21	15:57:31	Landsat 8	62.38916443	0.67	2.67	4.19		
79	21-Jul-21	16:03:45	Landsat 8	61.29353932	0.84	1.19	1.99		
80	30-Jul-21	15:57:39	Landsat 8	59.60748559	0.86	0.98	1.65		
81	06-Aug-21	16:03:53	Landsat 8	58.08933179	0.72	2.24	3.61		
82	14-Aug-21	15:04:00	Landsat 7	48.01591965	0.78	1.57	2.54	41.9840804	1.0129154
83	15-Aug-21	15:57:45	Landsat 8	55.89414317	0.84	1.26	2.1		
84	22-Aug-21	16:03:58	Landsat 8	54.01089576	0.53	3.66	5.59		
85	31-Aug-21	15:57:50	Landsat 8	51.38469741	0.77	1.75	2.84		
86	07-Sep-21	16:04:03	Landsat 8	49.19662115	0.79	1.62	2.66		
87	16-Sep-21	15:57:54	Landsat 8	46.23729378	0.79	1.58	2.6		
88	23-Sep-21	16:04:05	Landsat 8	43.84762947	0.77	1.66	2.68		
89	02-Oct-21	15:57:59	Landsat 8	40.7113483	0.74	1.99	3.22		

Appendix 2 : BayesVarSel Method

Figure S4.2. Correlation Matrix of the independent variables.



	LST	Prev_Inc	Per_Black	Per_immi
LST	1.00	0.00	0.00	0.00
Prev_Inc	0.00	1.00	0.34	0.46
Per_Black	0.00	0.34	1.00	0.42
Per_immi	0.00	0.46	0.42	1.00
Per_Vminority	0.00	0.58	0.52	0.89
Pop_Density_SQ_KM	0.00	0.46	-0.07	-0.12
Rate_Edu	-0.01	-0.23	-0.74	-0.59
Rate_inade	0.00	0.42	0.47	0.02
Rate_unaff	0.00	0.83	0.11	0.43
Rate_unsui	0.00	0.64	0.62	0.73
Unemp_Rate	0.00	0.64	0.63	0.69

	Per_Vminority	Pop_Density_SQ_KM	Rate_Edu
LST	0.00	0.00	-0.01
Prev_Inc	0.58	0.46	-0.23
Per_Black	0.52	-0.07	-0.74
Per_immi	0.89	-0.12	-0.59
Per_Vminority	1.00	-0.04	-0.54
Pop_Density_SQ_KM	-0.04	1.00	0.22
Rate_Edu	-0.54	0.22	1.00
Rate_inade	0.07	0.19	-0.41
Rate_unaff	0.45	0.41	-0.02
Rate_unsui	0.72	0.10	-0.62
Unemp_Rate	0.78	0.00	-0.57

	Rate_inade	Rate_unaff	Rate_unsui
LST	0.00	0.00	0.00
Prev_Inc	0.42	0.83	0.64
Per_Black	0.47	0.11	0.62
Per_immi	0.02	0.43	0.73
Per_Vminority	0.07	0.45	0.72
Pop_Density_SQ_KM	0.19	0.41	0.10
Rate_Edu	-0.41	-0.02	-0.62
Rate_inade	1.00	0.19	0.46
Rate_unaff	0.19	1.00	0.50
Rate_unsui	0.46	0.50	1.00
Unemp_Rate	0.41	0.40	0.78

	Unemp_Rate
LST	0.00
Prev_Inc	0.64
Per_Black	0.63
Per_immi	0.69
Per_Vminority	0.78
Pop_Density_SQ_KM	0.00
Rate_Edu	-0.57
Rate_inade	0.41
Rate_unaff	0.40
Rate_unsui	0.78
Unemp_Rate	1.00

n= 12460

P

	LST	Prev_Inc	Per_Black	Per_immi
LST		0.9038	0.8562	0.8739
Prev_Inc	0.9038		0.0000	0.0000
Per_Black	0.8562	0.0000		0.0000
Per_immi	0.8739	0.0000	0.0000	
Per_Vminority	0.7850	0.0000	0.0000	0.0000
Pop_Density_SQ_KM	0.8403	0.0000	0.0000	0.0000
Rate_Edu	0.2106	0.0000	0.0000	0.0000
Rate_inade	0.6462	0.0000	0.0000	0.0275
Rate_unaff	0.6159	0.0000	0.0000	0.0000
Rate_unsui	0.9310	0.0000	0.0000	0.0000
Unemp_Rate	0.7870	0.0000	0.0000	0.0000

	Per_Vminority	Pop_Density_SQ_KM	Rate_Edu
LST	0.7850	0.8403	0.2106
Prev_Inc	0.0000	0.0000	0.0000
Per_Black	0.0000	0.0000	0.0000
Per_immi	0.0000	0.0000	0.0000
Per_Vminority		0.0000	0.0000
Pop_Density_SQ_KM	0.0000		0.0000
Rate_Edu	0.0000	0.0000	
Rate_inade	0.0000	0.0000	0.0000
Rate_unaff	0.0000	0.0000	0.0383
Rate_unsui	0.0000	0.0000	0.0000
Unemp_Rate	0.0000	0.9896	0.0000

	Rate_inade	Rate_unaff	Rate_unsui
LST	0.6462	0.6159	0.9310
Prev_Inc	0.0000	0.0000	0.0000
Per_Black	0.0000	0.0000	0.0000
Per_immi	0.0275	0.0000	0.0000
Per_Vminority	0.0000	0.0000	0.0000
Pop_Density_SQ_KM	0.0000	0.0000	0.0000
Rate_Edu	0.0000	0.0383	0.0000
Rate_inade		0.0000	0.0000
Rate_unaff	0.0000		0.0000
Rate_unsui	0.0000	0.0000	
Unemp_Rate	0.0000	0.0000	0.0000

	Unemp_Rate
LST	0.7870
Prev_Inc	0.0000
Per_Black	0.0000
Per_immi	0.0000
Per_Vminority	0.0000
Pop_Density_SQ_KM	0.9896
Rate_Edu	0.0000
Rate_inade	0.0000
Rate_unaff	0.0000
Rate_unsui	0.0000
Unemp_Rate	

Table S4.2. Inclusion Probabilities from the BayesVarSel Method.

	Incl.prob.	HPM	MPM
LST	1	*	*
Prev_Inc	0.038		
Per_immi	1	*	*
Pop_Density_SQ_KM	0.0312		
Rate_Edu	1	*	*
Rate_inade	0.0318		

Appendix 3 : Spatiotemporal Model Specifications

Prior distributions were assigned for all model parameters. A vague uniform prior (0, 0001, 1000) was assigned independently to the standard deviations of the spatial, temporal and space-time random effect terms $\boldsymbol{\sigma_s, \sigma_u, \sigma_v, \sigma_\delta}$, respectively. The intercept $\propto$ has the improper uniform prior. The overall spatial random effect component was modelled using the Besag York Mollié (BYM model)[32]. The BYM model is a convulsion of a spatially structured random effect and a spatially unstructured random effect following a Gaussian distribution. The conditional autoregressive (CAR) prior with a spatial adjacency matrix W of Size $N \, x \, N$, where $w_{ij} = 0$, and the off-diagonal entries $w_{ij} = 1$, if areas *i* and *j* share common boundary lines and otherwise[33]. The CAR prior to the spatial random effects implies that adjacent neighbourhoods tend to have similar overall COVID-19 risks. The overall temporal component was modelled using the RW1 model by Fahrmeir and Lang (2001) to describe the overall time trend common to all neighbourhoods. The RW1 is a temporal adaptive process that represents a one-dimensional analogue of the ICAR, and the parameter at the first time point *t*, is assigned a vague/diffuse prior. A noninformative prior $Normal(0,0.0001)$ is assigned to the three regression coefficients $(beta1, beta2 \; and \; beta3)$.

In Model 2, a Type I space-time dependent structure was used where all parameters are similar to where they are in the "the space-time cube". In Model 3, a Type II space-time dependent structure was used where the temporal parameters in each neighbourhood *i*, are temporally smooth. The time trend pattern represented by the temporal parameters in area *j* ($i \neq j$) sis not assumed to be dependent similar to each other, even in the case *i* and *j* are spatially contiguous[31]. In Model 4, a Type III space-time dependent structure was used where the area-specific

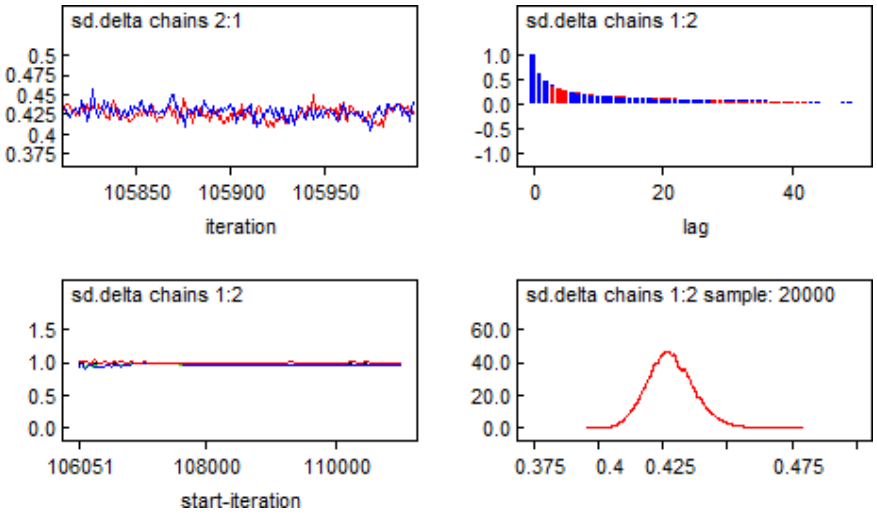
parameters are spatially smooth at each time point *t*. However, the spatial pattern represented by the area-specific parameters at time point *t* and the spatial pattern represented by the area-specific parameters at time point g ($t \neq g$) is not assumed to be dependent on or similar to each other, even in the case where *t* and *g* are close.

WINBUGS Code (Model 3: Space-time inseparable model with Type II interaction effect).

```
for (i in 1:N) {
for (t in 1:T) {
y[i,t] ~ dpois(mu[i,t]) # response variable Poisson likelikhood for the covid case count in #neighbourhoods i in time
t

mu[i,t] <- pop[i]*theta[i,t]  ## Population is assumed to be constant over time.

log(theta[i,t]) <-  alpha + beta1*EDU[i,t] + beta2*LST[i,t] + beta3*IMMI[i,t] + S[i] + U[i] + v[t] + delta [i,t]

RR[i,t] <- exp(delta[i,t])  ## Spatiotemporal Relative Risk

}  #close the for-loop over T
} #close the for-loop over N

for (i in 1:N) {
delta[i,1:T] ~ car.normal(tm.adj[],tm.weights[],tm.num[],prec.delta)
}

for (i in 1:N) {
U[i] ~ dnorm(0,prec.U)
}

S[1:N] ~ car.normal(sp.adj[], sp.weights[], sp.num[], prec.S)

v[1:T] ~ car.normal(tm.adj[],tm.weights[],tm.num[],prec.v)

# priors:
alpha  ~ dflat()
beta1 ~ dnorm(0, 0.00001)
beta2 ~ dnorm(0, 0.00001)
beta3 ~ dnorm(0, 0.00001)

### set prior on standard deviation scale
sigma.S ~ dunif(0.0001,10) ## Standard Deviation
sigma.U ~ dunif(0.0001,10) ## Standard Deviation
sigma.v ~ dunif(0.0001,10) ## Standard Deviation
sigma.delta ~ dunif(0.0001,10) ## Standard Deviation

### convert standard deviations to precision as required by WinBUGS

prec.S <- pow(sigma.S,-2)
prec.U <- pow(sigma.U,-2)
prec.v <- pow(sigma.v,-2)
prec.delta <- pow(sigma.delta,-2)

sd.S <- sd(S[1:N]) # unconditional SD of the spatially structured random effects
var.S <- pow(sd.S,2) # unconditional variance of the  spatially-structured random effects
sd.U <- sd(U[1:N]) # SD of the spatially-unstructured random effects
var.U <- pow(sd.U,2) # variance of the spatially-structured  random effects variance partition coefficient/spatial
fraction
sd.v <- sd(v[1:T])
```

```
var.v <- pow(sd.v,2) # variations of the temporal random effects
sd.delta  <- sd(delta[1:N,1:T])
var.delta <- pow(sd.delta, 2)
vpc.SP <- (var.S + var.U)/(var.S + var.U + var.v + var.delta)    ### VPC of spatial effects
vpc.TM <- var.v/(var.S + var.U + var.v + var.delta)    ### VPC of temporal effects
vpc.SPTM <- var.delta/(var.S + var.U + var.v + var.delta)    ### VPC of spatiotemporal effects

### the average COVID-19 rate per 1000 population per week in toronto
Toronto.average <- mean(theta[,])*1000

# Temporal Relative Risk

for (t in 1:T) {
temporal.RR[t] <- exp(v[t])
 }
# Spatial Relative Risk

for (i in 1:N) {
 spatial.RR[i] <- exp(S[i]+U[i])
 }
}
```

Figure S4.3. Trace plot (top left), autocorrelation plot (top right), and Gelman-Rubin plot (bottom left) for checking of convergence, and density plot of the posterior distribution (bottom right) of the standard deviation of spatio-temporal interaction effects (θ_{it})



Appendix 4: Sensitivity Analysis Results

Table S4.3. Comparison of the DIC outputs from the sensitivity test.

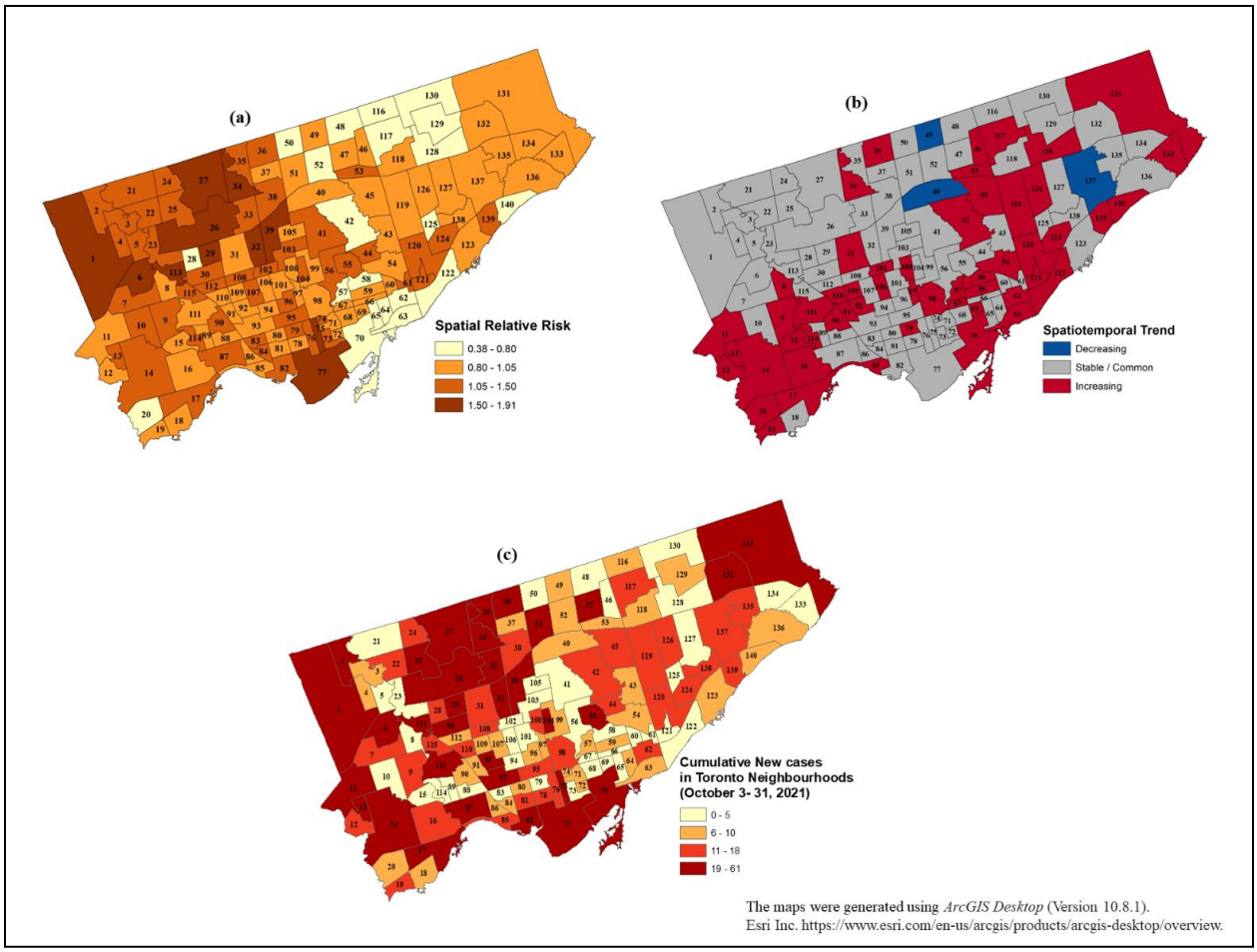
	Dbar	Dhat	pD	DIC
Model 3 with Uniform Prior (0.01,10)	52247.4	49015.6	3231.73	55479.1
Sensitivity Test				
Model with Gamma Prior (0.005,0.005)	52263	49014.2	3249.63	55513
Model with Uniform Prior (0.01,1000)	52265.6	49013.2	3252.41	55518

Table S4.4. Comparison of the parameter estimates from the sensitivity test.

Parameters	Model 3 with Uniform Prior (0.01,10)	Model with Gamma Prior (0.005,0.005)	Model with Uniform Prior (0.01,1000)
alpha	-8.556	-8.557	-8.556
beta1	-0.3258	-0.3283	-0.3282
beta2	0.01516	0.0148	0.01552
beta3	0.08472	0.08279	0.08244
var.S	0.04108	0.05567	0.0573
var.U	0.05079	0.0368	0.03744
var.delta	0.1834	0.1833	0.1838
var.v	3.481	3.454	3.48
sd.S	0.2001	0.2345	0.2381
sd.U	0.2237	0.1891	0.1909
sd.delta	0.4282	0.428	0.4286
sd.v	1.865	1.858	1.865

Appendix 5: Map validation

Figure S4.4: Comparison of the resulting Spatiotemporal trends and risk maps with the 4-week post cumulative new cases. (a) map of spatiotemporal trend, (b) spatial risk derived from the Bayesian models, (c) cumulative number of new cases during the post-study period.



Appendix B: (Chapter 5)

Fig S5.1. Frequency distribution of the incidence rate/1000 persons in the neighbourhoods of Toronto.

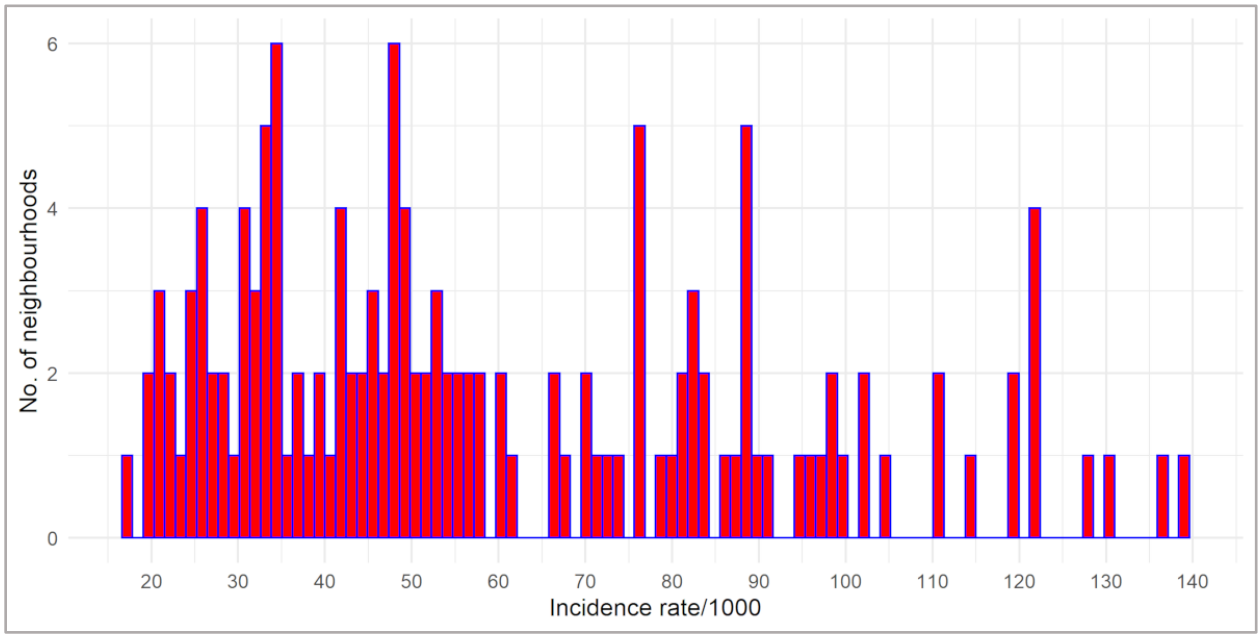
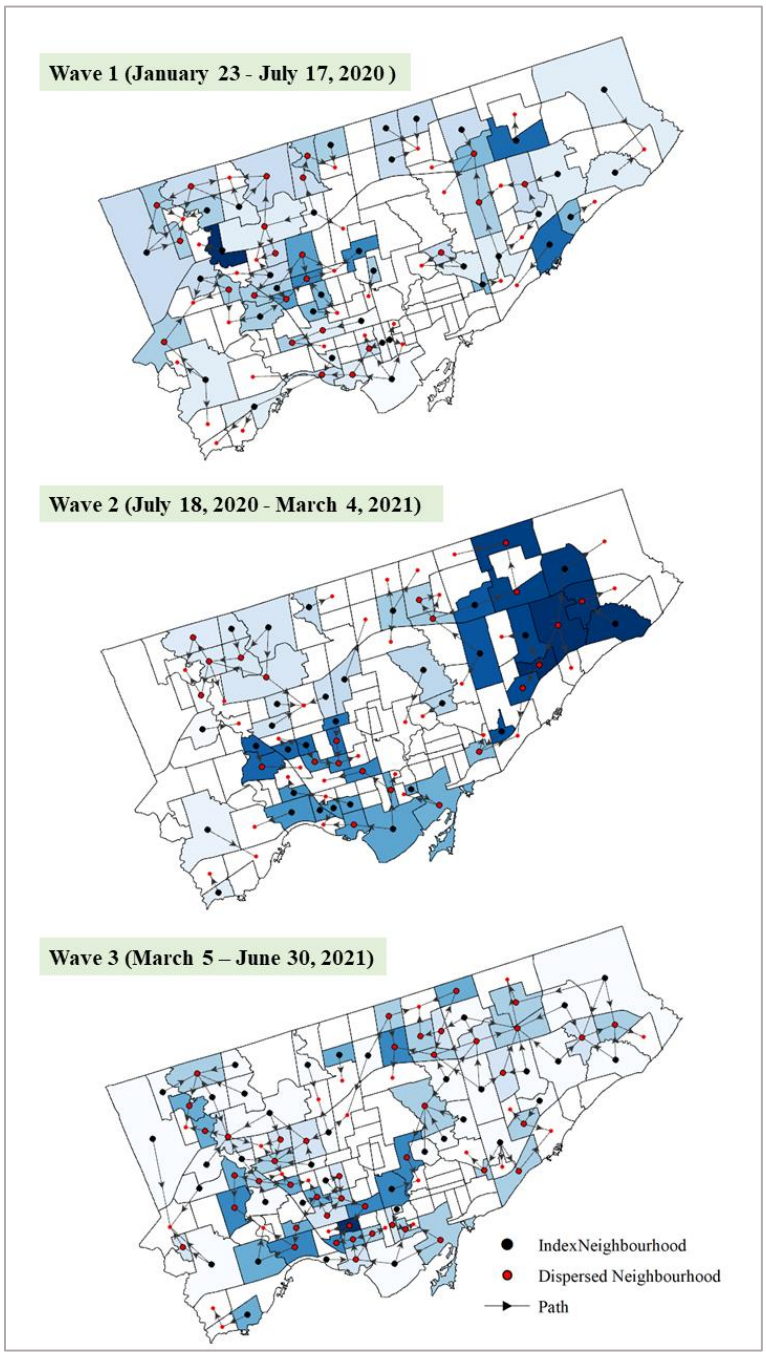
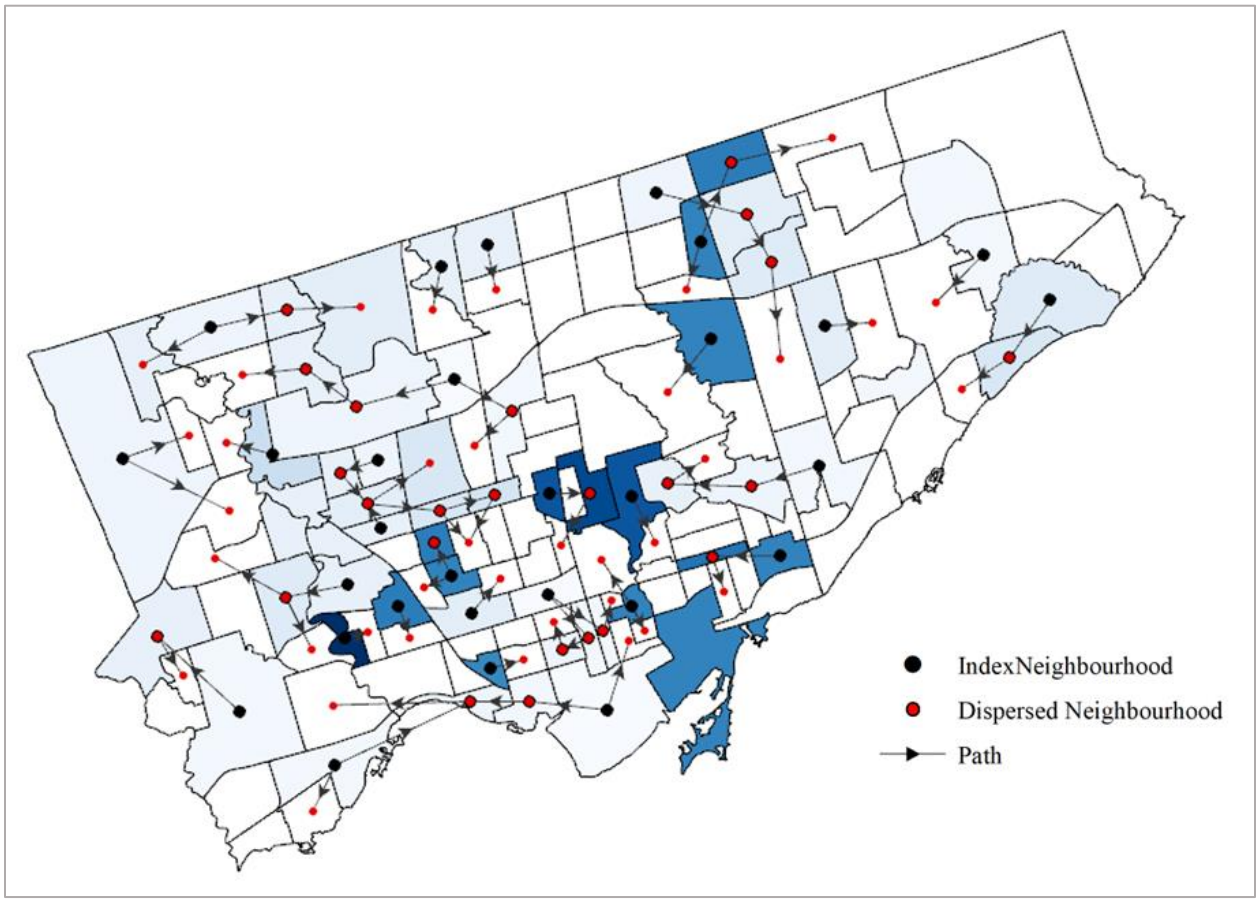


Fig S5.2. Spatiotemporal spreads of COVID-19 during the first three COVID-19 infection waves in Toronto.



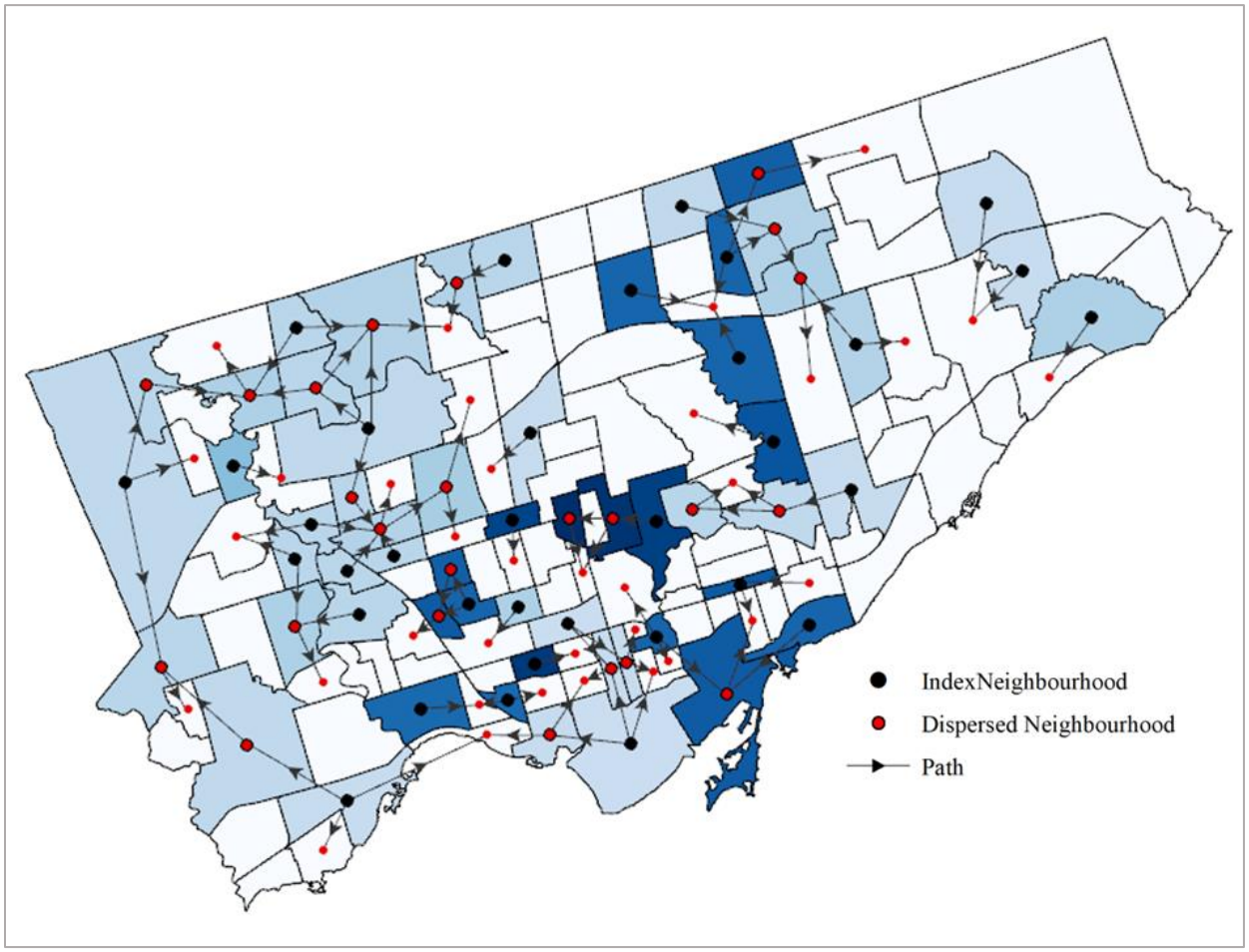
Note: The lighter the colour the earlier the onset of the first outbreak in the neighbourhood.

Fig S5.3. Spatiotemporal spreads of COVID-19 with a single index (source) neighbourhood using nearest distance method in Toronto, January 2020 - July 2021.



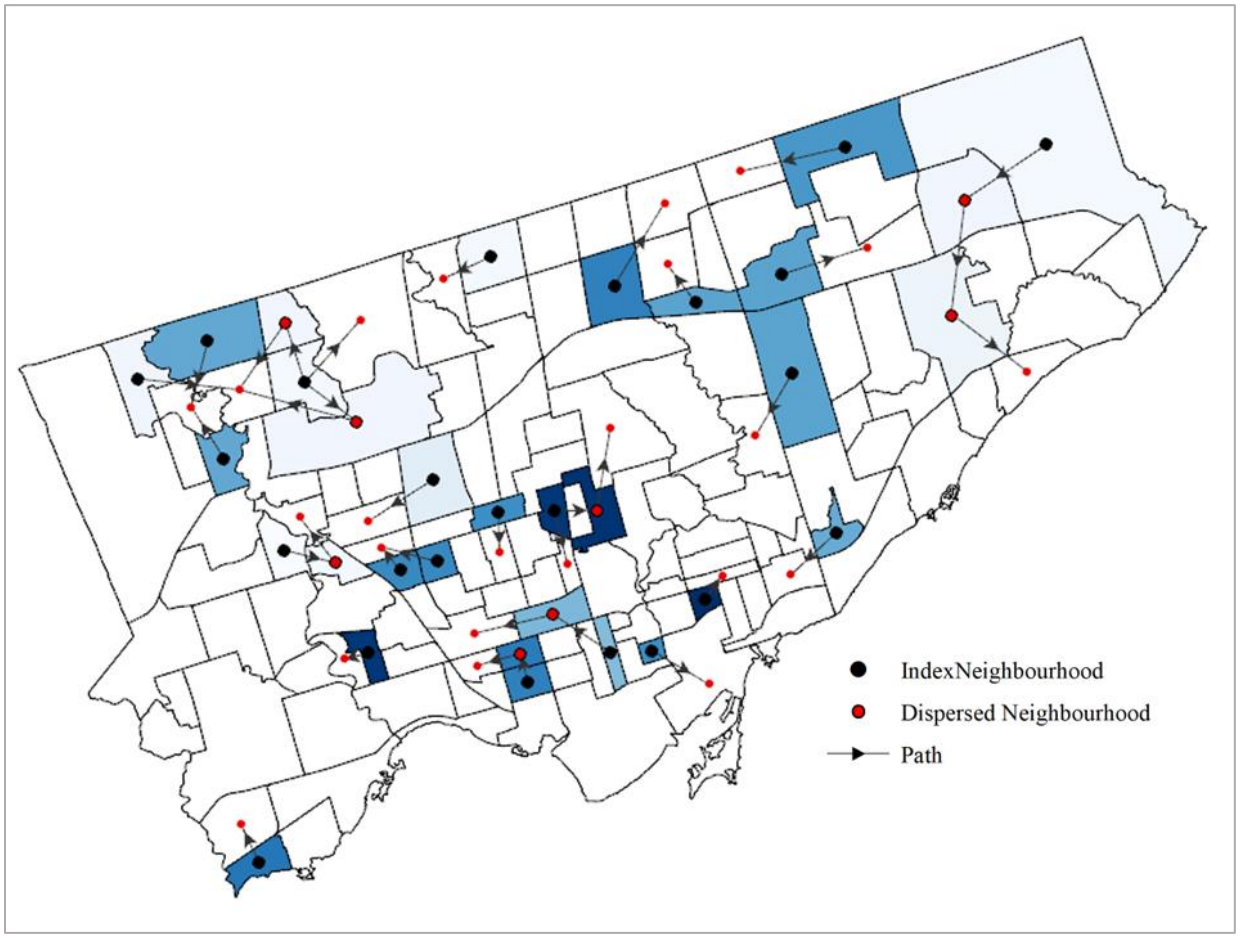
Note: The lighter the colour the earlier the onset of the first outbreak in the neighbourhood.

Fig S5.4. Spatiotemporal spreads of COVID-19 using the median to define the outbreaks in Toronto, January 2020 - July 2021.



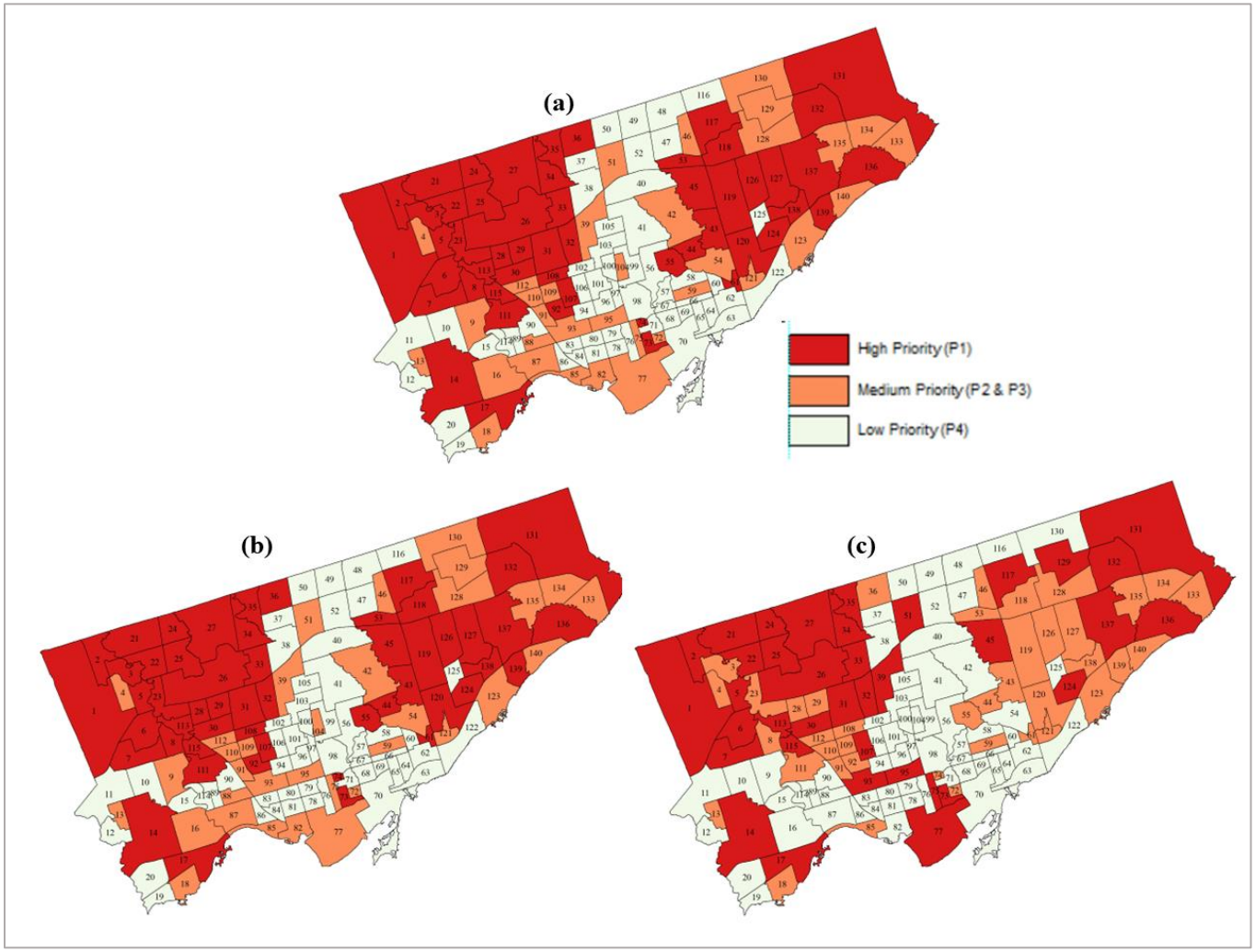
Note: The lighter the colour the earlier the onset of the first outbreak in the neighbourhood.

Fig S5.5. Spatiotemporal spreads of COVID-19 using one week gap between the index and dispersed neighbourhoods.



Note: The lighter the colour the earlier the onset of the first outbreak in the neighbourhood.

Fig S5.6. Hotspot maps using the epidemiological parameters.

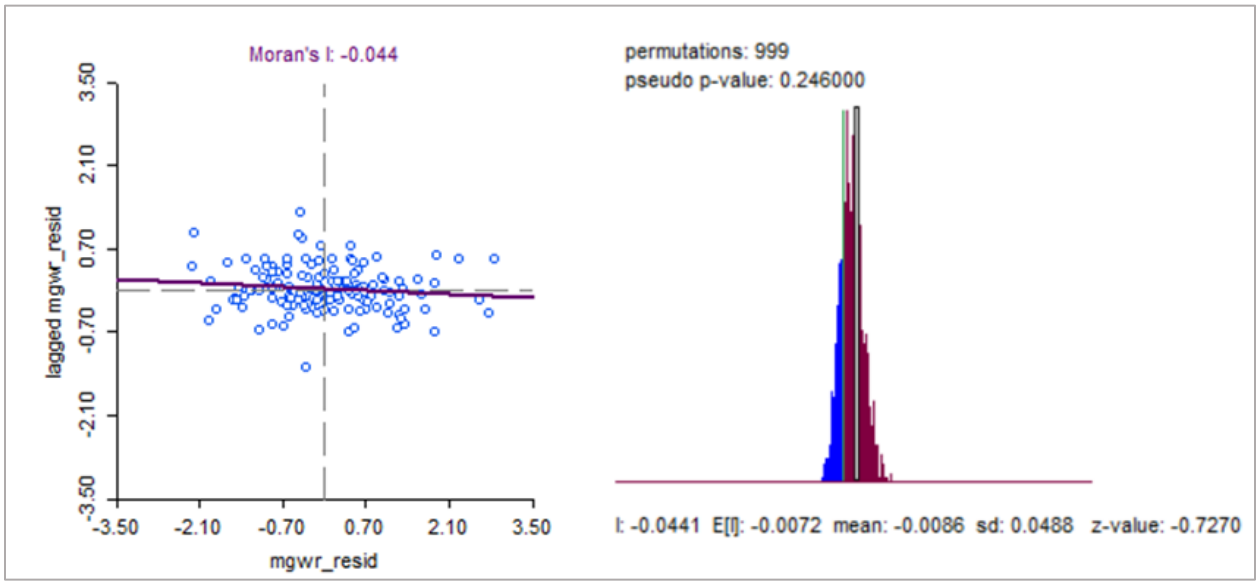


a) 50th percentile values for both incidence rate and persistence (50 cases/1000 persons, 50 weeks)
b) 70th percentile value (76.29 cases/1000 persons) for incidence rate and 50th percentile value (61 weeks) for persistence
c) 50th percentile value for incidence rate (50.37 cases/1000 persons) and 70th percentile (65 weeks) persistence.

Note : The numbers in the chart and in the map are the neighbourhood IDs.

Appendix C: (Chapter 6)

Fig S6.1 Assessment of spatial patterns of the residuals obtained from the M(GWR) model.



Global Moran's *I* was performed to test for the presence of spatial autocorrelation among the residuals obtained from the M(GWR) model (Chapter 3). GeoDa (v1.18.0) was used to perform this analysis. For the Global Moran's *I* statistic, the null hypothesis states that the residuals being analyzed are randomly distributed among the geographic units in the study area, and the spatial processes promoting the observed pattern of values are random in space.

The results of the Global Moran's *I* test are presented in Figure S6.1. In the left figure, the residuals are on the x-axis, and their spatially lagged counterparts are displayed on the y-axis. The slope of this line corresponded to Moran's *I* and its value (-0.044). A negative Moran's *I* value implies that the residuals do not show the presence of clustering and that the data is dispersed. A randomization with 999 permutations (generally sufficient for yielding a reliable inference) was selected, and a histogram was generated (right figure). The result of the pseudo-p-value (0.246) rejects the alternative hypothesis implying that the residuals of the MGWR model are not spatially clustered. Therefore, we can conclude

that the assumptions of normality of the errors of the M(GWR) model were met, and the assumption of the homoscedasticity of the error residuals was not violated.