Spatiotemporal Analysis of Human Mobility based on Land Use Types in the Greater Toronto Area during COVID-19 Pandemic

by

Yiqing Wu

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Geography

Waterloo, Ontario, Canada, 2023

Author's Declaration

I hereby declare that I am the sole author of this 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.

Abstract

The 2019 Coronavirus disease COVID-19 is an infectious respiratory disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It undoubtedly poses a huge challenge in terms of public health and social impact worldwide. The Ontario government implemented a series of non-pharmaceutical interventions (NPIs) prior to vaccination to prevent large-scale outbreaks in the Great Toronto Area (GTA), which is the most densely populated region in Ontario. Detecting and analyzing human mobility during the pandemic can help decision makers assess the effectiveness of policy implementation, in order to better respond to similar events in the future.

Geotagged Twitter data serves as an important source of volunteered geographic information (VGI). Anonymized geotagged tweet in the GTA in 2020 using the Twitter Academic API are used to analyze inner-city human mobility. The results provide a longer-term insight into how human activity is affected by the pandemic as well as government orders.

In this thesis, human mobility spatiotemporal patterns in the GTA are found to be close to patterns founded in the previous studies. People are affected more by the severeness of the first outbreak. More people stay at home rather than in commercial areas, schools, and workplaces. Human mobility in open spaces is affected by seasons besides policy effects. Human mobility in utility and transportation areas is related to the properties of the areas they connect. Most of the policies received significant reflections within one week of release, but milder policies resulted in insignificant human mobility changes. Human mobility patterns in most land use types have moderate correlation with the Google Community Mobility Report. Even so, some limitations still exist.

Acknowledgements

First of all, I would like to give my thanks to all those who have helped me in this thesis.

My sincere gratitude and appreciations go firstly to my supervisor, Professor Jonathan Li for his continuous support and encouragement over the year. His profound knowledge and insightful guidance make me understand geographical information more deeply and inspires me a lot in my thesis. It is a great honour to study under his supervision.

I would like to express my sincere thanks to my thesis committee members, Dr. Michael Chapman, Professor at the Toronto Metropolitan University and Dr. Zahid Butt, Assistant Professor at the University of Waterloo. I am also grateful to Dr. Sabrina Li, Assistant Professor at the University of Nottingham. Thanks for their valuable comments and suggestions, which help me a lot to improve the quality of the thesis.

My appreciation also to Geospatial Intelligence and Mapping Lab members for their professional and warm support. A special thanks to my graduate fellows and my friends Sarah Narges Fatholahi, Charlotte Zhuoran Pan, Jingtian Tan, Yikai Yao, and Siyu Li for their vitality and encouragement. Thanks for their company during my graduate studies.

Last but not least, I would like to express my hearty thanks to my parents and families for their unconditional support and endless love throughout my life.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CMA	Census Metropolitan Area
COVID-19	Coronavirus Disease
GDP	Gross Domestic Product
GI	Geography Information
GTA	Greater Toronto Area
JSON	JavaScript Object Notation
NPI	Non-Pharmaceutical Intervention
PHEIC	Public Health Emergency of International Concern
POI	Point of Interest
QEW	Queen Elizabeth Way
SARS	Severe Acute Respiratory Syndrome
VGI	Volunteered Geographic Information
WHO	World Health Organization

Chapter 1

Introduction

1.1 Problem Statement

COVID-19, the Coronavirus disease 2019 is a contagious respiratory disease caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). The disease is first reported in late 2019 and quickly broke out in most countries and territories until March 2020. Thus, World Health Organization (WHO) accessed COVID-19 as a pandemic. The COVID-19 pandemic brings not only health risks, but also significant and ongoing impacts on global economic and social activities. By the end of 2022, the COVID-19 pandemic caused more than 732 million confirmed cases and more than 6.7 million death cases. Among them, there are approximate 4.5 million confirmed cases and 49,000 death cases in Canada (World Health Organization, 2020). Industries dependent on trade, labor, and global industrial chains suffered severe damage during the pandemic (Wei et al., 2021). People's lifestyles have undergone structural changes; therefore, consumer behaviors changed according to their psychology, income, market supply and demand (Habibi et al., 2022). Almost all businesses in Canada were affected by the shutdown in Spring 2020, until the end of 2020, overall economic activity is still lower than the pre-pandemic levels. The economic downturn caused a rise in unemployment. The loss of jobs and working hours has affected more than one million workers in Canada (Statistics Canada, 2021). The COVID-19 pandemic also had negative impacts on people's mental health in a variety of ways, such as reduced contact with family and friends (Okan et al., 2021), changes in demographic characteristics (Lei et al., 2020), vulnerability of special groups (Twenge & Joiner, 2020), etc. In addition, crime opportunities and observations changed with pandemic and policy implementation, which calls for law enforcement to update its focus and resource allocation to different types of crime (Díaz et al., 2022; Abrams, 2021). Overall crime level in Canada decreased during the pandemic but some types of crime increased due to lack of sightings. Hate crimes peaked in 2020, reflecting the discrimination caused by the COVID-19 pandemic (Statistics Canada, 2022).

The challenge is to reduce the spread of the COVID-19 pandemic and its negative effects while safeguarding the productive lives of people as much as possible. Before vaccine became widely available, non-pharmaceutical interventions (NPIs) were implemented in various countries and regions to prevent large outbreaks of the COVID-19 pandemic, including but not limited to masking, social distancing and limiting the number of people gathered at the individual level, to city-scale lockdowns or curfews, to nationwide border closures. Those NPIs have been shown to significantly reduce the growth rates in Spain (Orea & Álvarez, 2022), the USA (Li et al., 2021), Italy (Bertuzzo et al., 2020), Japan (Yabe et al., 2022), China (Fang et al., 2020) and many other countries. However, the measure and its strictness of implementation can vary tremendously from country to country. A typical example is the Zero-COVID strategy implemented in many parts of China, where residents are banned from leaving their communities if there is any confirmed case within the community. This measure protected more healthy people from the potential risk of infection, but its strict enforcement over a long period has also led to mood swings and protests by residents (Bai et al., 2022). Most Western countries, on the other hand, have not implemented such long and severe lockdown strategies, which has led to a continuous presence of confirmed numbers and risk of close contact in these areas. Through a review the literature of health impact by NPIs during the pandemic, although most studies focused on the direct health impact of the COVID-19 pandemic, a small number of studies focused on the social impact and economic harm that NPIs may cause in terms of reduced life happiness and lack of social welfare (Chiesa et al., 2021). Therefore, the stringency of policy implementation should balance effectiveness and social impact.

The Greater Toronto Area (GTA) is an urban cluster consisting of 25 cities in four surrounding regions, with the City of Toronto as its centre. It is the most densely populated area in Ontario as well as in Canada, also contributes significantly to the financial and manufacturing sectors of North America. On the one hand, the GTA presents cultural diversity and a high level of inclusiveness towards new immigrants, with the benefit that international immigration has contributed to the transition to a post-industrial phase and to the development of a knowledge-based economy. On the other hand, diversity brings more vulnerabilities to pandemics. As early as 2003, Toronto was the only non Asian city to be severely affected by the SARS (Brail & Kleinman, 2022). During the pandemic, the infection and mortality rate in Toronto was disproportionate compared to its population (Crawley, 2020). In studies of US states and England and Wales during the pandemic, significant differences in mortality and infection rates are observed between races and communities. Differences between cultures, labor market, housing condition, and other diverse characteristics may cause those bias (Polyakova et al., 2021; Nathan, 2021). Connectivity and diversity between municipalities cause more difficulties and challenges to the response to the COVID-19 pandemic in the GTA. Therefore, the research on the GTA can provide some suggestions of urban clusters dealing with crisis events in future.

Human mobility patterns present the characteristics of people moving through space. It has a high application value in several fields. In epidemiology, it is common to monitor human mobility to predict models of infection and assess the effectiveness of policies (Santamaria et al., 2020; Chinazzi et al., 2020; Rahman & Thill, 2022). Most NPIs developed for respiratory infectious diseases such as the COVID-19 pandemic are designed to reduce human-to-human contact to cut off the route of infection. Therefore, NPIs with this goal can use human mobility change as a criterion for assessing whether the policy is effective. Human mobility data can be accessed through social media (Bisanzio et al., 2020), mobile network operators (Yabe et al., 2022), public transport travel records such as bike-sharing (Song et al., 2022), metro and flight information and so on (Sy et al., 2021; Li et al., 2021). However, from existing research, when mobility data with distinct origin and destination locations are used to analyze changes in mobility, travel distance and travel time are usually extracted from the raw data to calculate changes in mobility index, with origin-distance matrices and radius of gyration being the more commonly used methods. These approaches allow different distance thresholds to be set for a wider range of spatial scales, but it eliminates some of the spatial detail, such as the types of places people visit, when extracting movement distances. In contrast, in some cases exploring 'where people go', researchers usually extract the coordinates of user or their posts directly for spatiotemporal analysis. For example, Jiang et al. (2021) used geotagged tweets to track spatiotemporal variation in human mobility on land use polygons within New York City, also validated against the categories provided by the Google Community Mobility Report. Heo et al. (2020) used the most common locations of Facebook users to investigate changes in human mobility within parks and forests in Maryland and California, which identifying green spaces as providing good options for keeping social distances and outdoor activities during the pandemic. Both Jiang et al. (2021) and Heo et al. (2020) used the number of active users within a given area to represent human mobility.

In the current study, there are a small number of applications of human mobility analysis during pandemics in Canada, an example is that Klar (2022) used mobile device data provided by Telus Insight to represent human mobility in Ontario by calculating the radius of gyration. There are no published studies that have applied human mobility analysis at the sub-city level or used social media to capture human mobility in Canada. Therefore, this study will focus on monitoring spatial and temporal variation in human mobility based on different land use types during the COVID-19 pandemic and use this to assess the effectiveness of NPIs in the GTA.

1.2 Research Questions and Objectives

The aim of this study is to address the following questions:

- How does each wave of the COVID-19 pandemic and corresponding policies affect the spatial patterns of human mobility within the GTA?
- Do the types of destinations people choose to visit change over time?
- What are the differences between the patterns of human mobility detected by Twitter and the community mobility reports provided by Google?

This study enriches the analysis of spatial and temporal variation in human mobility in terms of 'where people go' and informs the use of voluntary geographical information (VGI) to assess the effectiveness of GTA policy implementation during pandemics. The goal of this study is to explore the spatial and temporal variation in human mobility in the GTA during the COVID-19 pandemic particularly in 2020 based on different land use types. To achieve this goal, the following objectives are designed:

- To explore spatial pattern variation of human mobility within the GTA in both monthly and weekly during the pandemic based on land use polygons.
- To determine if there is a correlation between this pattern of change and the number of confirmed cases and measures issued by the Ontario government.
- To explore the temporal change pattern of human mobility based on land use type and validate using the Google Community Mobility Report.

1.3 Thesis Structure

This thesis consists of five chapters.

Chapter 1 describes the problems and motivations of the study, follow by the research questions and objectives as well as the structure of the study.

Chapter 2 reviews the existing literature on the use of changes in human mobility to assess policy effectiveness and briefly reviews the methods used to assess policy effectiveness during the COVID-19 pandemic.

Chapter 3 describes the study area (the GTA), the dataset used, and detailed spatiotemporal analysis methods.

Chapter 4 presents the results of spatiotemporal analysis and provides a discussion of the findings.

Chapter 5 concludes the key findings and identifies some of the limitations of this study, also provides recommendations for future research.

Chapter 2

Background and Related Studies

This chapter reviews existing literature and focuses on research work related to the application of Volunteered Geographical Information (VGI) to crisis response, especially in epidemiology. Section 2.1 discusses the value of VGI. Section 2.2 focuses on the application of human mobility in epidemiology. Section 2.3 discusses some research gaps. Section 2.4 provides a summary of this chapter.

2.1 Value of Volunteered Geographical Information

Volunteered Geographical Information can be defined as geographic information that is voluntarily created, collected, and disseminated by individuals using digital technologies (Goodchild, 2007). Compared to traditional labor- and financial-intensive data collection methods, encouraging citizens to observe the world as sensors allows for timely data collection with experiential aspects (Ferster et al., 2018). This is benefit from citizens who are more familiar with the local environment and more sensitive to what is happening there. Theoretically, every citizen can synthesize and interpret local information intelligently (Goodchild, 2007). As a result, the potential of VGI for applications within multiple domains is receiving increasing attention from social scientists, especially in crisis response (Haworth et al., 2016; Haworth B. T., 2018; Hicks et al., 2019; Joshi et al., 2020), smart cities (Basiouka et al., 2015; Mozas-Calvache, 2016; Attard et al., 2016), and biodiversity studies (Jacobs & Zipf, 2017; Brown et al., 2018). Goranson et al. (2013) indicated that VGI has contribute to public health through allowing users to provide detailed location-based information to fit customized study areas. In epidemiology, VGI allows users to track their historical locations, which is helpful to identify weak spatial relationships. The local tapestry carried by VGI can also use to reflect users' opinions in food services or new policies.

The value of geographic information (GI) is usually measured by assessing its economic value through quantitative methods (Longhorn & Blakemore, 2008), such as cost-benefit analysis (Hall et al., 2000) or value chain methods (Mancini, 2013), also, analyzing its social benefits through qualitative fields. It is much more complicated to qualify the social value of GI than quantifying the economic value. The social value of GI can be evaluated by whether services are improved, decision-making is more informed, and there is an increase in the ability of vulnerable groups to use public geographic information and spatial technologies (Feick & Roche, 2012). VGI is applied in some commercial software as a supplement to private authoritative GI to reduce data collection cost. For example, the reporting problem function in Google Maps effectively reduces the cost of data maintenance. Besides, Open Street Map, as a mature VGI project, also provides a stable economic basis for the company to broaden its business scope. In addition to the economic value of VGI mentioned above, the social value of VGI comes to highlight in some special cases. After the Haiti earthquake, when the map resources available in the affected areas are insufficient to address the demands of rescue and resource allocation, VGI is the only way that can be updated timely in a crisis scenario. Its social and economic value cannot be compared with private authoritative GI in the traditional way (Roche et al., 2013). In addition, citizens by participating in VGI activities are able to increase their awareness, expression and use of geographic information and location technologies, as well as improve their spatiality by activate spatial skills. This is a great improvement to the spatial support proposed by Williamson et al (2010).

While VGI has demonstrated distinct economic and social value in many areas, especially in crisis management, there are still some negative aspects. As a relatively new source of GI, the quality of data is a major challenge of VGI. During a disaster, people may send a large number of messages to request or donation. However, filtering inappropriate or unnecessary information can be costly, which may block emergency services and affect the resource allocation for emergency response (Holguín-Veras et al., 2014). Therefore, integrating the information provided by informal volunteers requires much carefulness to avoid weakening the adaptability and response capacity contribute to emergency management (Whittaker et al., 2015). In addition, social medias serve as one of the primary sources of VGI. The spread of fake news can affect community disaster resilience (Haworth et al., 2018). Vosoughi et al. (2018) found that the false news always spread faster and farther than the real news, and the bots spread the real news and the false news at the same rate. This means that real users are more likely to accelerate the spread of fake news.

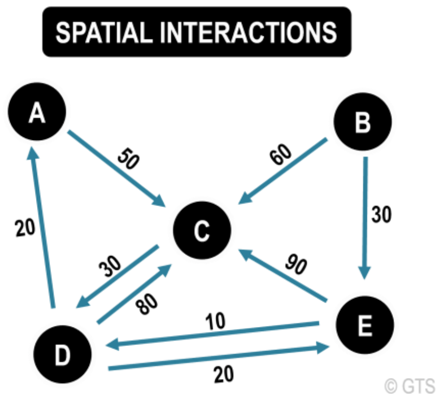
Overall, VGI still has great potential in crisis management as the technology evolves and algorithms update. The next section will focus on reviewing the usage of VGI to calculate human mobility in epidemiology.

2.2 Human Mobility in Public Health

The social value of VGI is not only shown in crisis management, such as natural disasters and man-made incidents, but also in public health. The main data sources of VGI are social media, cell phone location information, public transportation statistics and so on (Hu et al., 2021). Social media, in particular, can record not only people's aspects but also the location of users near real time. This makes social media as a major data source for many studies in public health. Social media data are widely used in studies of (1) infectious diseases such as H1N1 (Signorini et al., 2011), Ebola (Hossain et al., 2016), Zika (Abouzahra & Tan, 2021), Dengue (Souza et al., 2019), COVID-19 (Liu et al., 2022), etc., (2) non-infectious diseases such as obesity (Waring et al., 2018), pollen allergies (Gesualdo et al., 2015), etc., and (3) mental health such as suicidal tendencies (McClellan et al., 2017) and so on.

Human mobility is an important element in monitoring disease spreading in epidemiological studies, especially for human-mediated diseases. It is commonly used in studies of disease spreading model prediction (Zheng et al., 2021; Hu et al., 2021) and policy effectiveness evaluation (Xiong et al., 2020; Nguyen et al., 2021). Assessing the effectiveness of policies is important when responding to a pandemic. Under-react policies are not effective in limiting the spread of disease, and the increasing case numbers will bring the health care system under enormous pressure. On the other hand, overly strict policies are unhelpful to the proper functioning of society and the economy, also tend to create negative emotions among the population under strict control. Lai et al. (2020) generated a model of susceptible-exposed-infectious-removed (SEIR) based on travel network and found that the case numbers in China would increase exponentially without the influence of NPIs. Also, an earlier implementation of the policy would have greatly reduced the number of people and cities affected by the COVID-19 pandemic. In addition, NPI during the COVID-19 pandemic has the potential to cut off transmission routes for other infectious diseases. Wu et al. (2022) studied how the NPI implemented in Guangzhou during the COVID-19 pandemic affected the transmission of hand, foot, and mouth disease and found that there was an increase in hand, foot, and mouth disease cases after the relaxation of the NPI for the COVID-19 pandemic. This indicated that the effectiveness of NPI was not only done for the current epidemic but also for the prevention of potential epidemics.

The main data sources for acquiring human mobility are geotagged tweets and cell phone location data. Mobility reports from Google and Apple are sometimes used as a standard to validate the results. There are three main approaches to calculating human mobility; the first is to calculate the frequency of people's movement, such as the OD matrix. The second is to calculate the size of people movement through the radius of gyration (ROG) and travel distance, and the third is to represent human mobility using active user numbers. OD matrix and the radius of gyration are two primary methods use cellphone location data. The principle of the OD matrix is to calculate the directional flow between a series of positions, which represent the frequency of people move from a given origin to a given destination. Figure 2.1 shows the most basic OD matrix generation method (Rodrigue et al., 2006).



O/D MATRIX

	A	B	C	D	E	Ti
A	0	0	50	0	0	50
B	0	0	60	0	30	90
C	0	0	0	30	0	30
D	20	0	80	0	20	120
E	0	0	90	10	0	100
Tj	20	0	280	40	50	**390**

Figure 2.1 Schematic of Origin-Destination Matrix (Source: Rodrigue et al., 2006).

Static OD flows are estimated by three main methods, maximum likelihood (Cascetta & Nguyen, 1988), generalized least squares (Cascetta, 1984), and Bayesian (Maher, 1983). Dynamic OD is divided into n unit times over a period of time using n vectors of link counts to estimate n o-d matrices (Cascetta et al., 1993).

Travel distance and the radius of gyration indicates the distance and characteristic distance a person travels over a period of time, respectively. Liu et al. (2018) calculated the human radius of gyration (r_g) using cell phone movement data by the following equation:

$$r_g = \sqrt{\frac{1}{N}\sum_{i=1}^{N} n_i(\vec{r_i} - \overrightarrow{r_{cm}})^2} \tag{2.1}$$

where N is the total number of places, n_i means the visit frequency or spending time at the ith location, $\vec{r_i}$ is the geographic coordinate for the ith location, and the $\overrightarrow{r_{cm}}$ represents the coordinates of the weighted trajectory center of mass. These two approaches prefer a dataset that records origins and destinations, usually using cellphone location data. Some studies have also used tweet data for the calculation, but it is necessary to extract users who tweeted multiple times in a unit time. The amount of data required is larger, so it is often used for studies on a city scale or larger. Also, some details may be lost during them, such as what locations humans visited at what times.

During a pandemic, understanding the purpose of people's travel is an important factor in risk prediction. As a simple example, a person who travels to a store or a park within 5 km has a different risk of infection. Therefore, focusing on where people visit can theoretically lead to a better understanding of the risk of infection. Policy adjustments can also be made to encourage economic recovery in the post-epidemic era based on the types of hot destinations people visit. In addition to those two measures of human mobility based primarily on cell phone mobile data, a subset of studies using social media data to monitor human mobility though another qualitative approach based on point of interest (POI).

This approach used to classify social media check-in data by specific time periods and destination types, depending on the temporal and spatial scale of the study interested in. The number of active users at each epoch is used as a representative of human mobility. The amount of change is derived as a percentage using the difference between the observed value and the baseline value, divided by the baseline value. This method is used in several studies focus on close to or smaller than city scale (Yang et al., 2019; Jiang et al., 2021; Heo et al., 2020 & Chen et al., 2019). Yang et al. (2019) used geo-tagged Sina Weibo messages to discover patterns of human mobility in the Wuhan China Address University community. Jiang et al. (2021) used geotagged tweets to explore changes in human mobility based on land use type in New York City during the pandemic. Check-in data from an app based on Foursquare called Swarm was used to analyze the spatial and temporal variation of human mobility in Melbourne (Singh et al., 2018). Heo et al. (2020) studied in green spaces within Maryland and California state also using the similar method because a large green space can across more than one city. These studies achieve good results in using geotagged social media data to explore inner-city human mobility changes while focusing on destination types, as well as verify this qualitative approach is also effective to analysis human mobility change in a finer spatial granularity.

2.3 Gaps in Research

Most of the existing studies using the first two methods apply in municipal-scale to global-scale studies. Most of them focus on how the range of human activity impacted by the COVID-19 pandemic, such as people reduced inter-city or long-distance travel in response to the COVID-19 pandemic. These methods have been developed over time to obtain more realistic results in representing human mobility, which require higher data quality and a large amount of data, also preferably collect data using certain time intervals. There are a few studies that have used mobile network data to explore human mobility changes in Canada and Ontario, such as Klar (2022) used cellphone location data provided by Telus to calculate ROG and travel time to analyze human mobility changes in Ontario during a pandemic. Xue, et al. (2021) used Google Community Mobility to assess the impact of NPI on the spread of the COVID-19 pandemic within nine Canadian provinces. Dainton and Hay (2021) also used Google Community Mobility report to explore the relationship between polices, human mobility, and the spread of the COVID-19 pandemic in the GTA. Google Community Mobility Report is an aggregated dataset with a low spatial resolution that only reaches the regional scale, which makes it difficult to study in municipal human mobility analysis. However, these studies focus less on the purpose of people's travel, their spatial scale is also relatively larger. Some necessary intercity travels may not be reduced by the pandemic, and the destinations people travel may be influenced by the pandemic or policies.

Existing studies barely produced spatiotemporal analysis of human mobility below the city scale in Canada, but there are some previous studies using geotagged social media data for qualitative human mobility analysis in other cities such as Wuhan and New York. Those studies explained human mobility patterns in a finer spatial scale with a higher flexibility in study areas and study periods. Using the number of check-in data from Sina Weibo in The Chinese University of Geosciences Wuhan (CUG Wuhan) community, Yang et al. (2019) found that human mobility within the community was significantly influenced by vacation and gender. This study shows human mobility among university students in general and explores the potential of using the number of check-in data as a proxy for human mobility. Jiang et al. (2021) used geotagged tweets to explore human mobility changes in New York City during the early stages of the pandemic, which took advantages of the timeliness of geotagged tweet data. However, due to the short study period between February 16 and May 30, 2020, it is not available to explore the human mobility change patterns during the recovery phase and follow-up outbreaks. In addition, only one representative policy is mentioned in their study, which is the stay-at-home order issued in New York on March 22, 2020. Therefore, the relationship between policies and human mobility over a longer period of time is still a topic worthy of study.

To date, most studies applied in Canada use the Google Community Mobility Report directly as a proxy for human mobility at large spatial scales. Also, there is no precedent using geo-tagged tweet counts to explore human mobility changes at suburban scales for the time being. This thesis plans to draw on the approach proposed by Jiang et al. (2021), uses the number of geotagged tweets to represent human mobility changes based on land use polygons within the GTA over a longer period during the COVID-19 pandemic, which is over two outbreaks in 10 months from March to December in total, as well as studies in the impact on intra-city human mobility caused by the strictness of policies over a longer period.

2.4 Chapter Summary

In this chapter, the value of VGI and the applications of human mobility in public health are recognised. Meanwhile, three basic approaches to defining human mobility from existing research are discussed. Finally, some research gaps are indicted.

In the field of public health, the usage of VGI to explore human mobility patterns is becoming a general approach. Although VGI data sources have some biases, they are undeniably more efficient and flexible than community-based surveys and other traditional methods. The social medias, as one type of the main data sources for VGI, have the advantages of VGI but also some limitations. Geotagged social media data represents only a small fraction of all social media data. However, when the base is large enough, geotagged data is equivalent to a sub-sample that can represent almost the entire population. This makes social media data to be an emerging data source due to its easy accessibility and less privacy concerns.

Human mobility is primarily used in pandemics to model disease transmission and analyze the effectiveness of policies. The existing literatures indicate that the implementation of NPIs can be effective in reducing the spread of disease. However, the strictness of NPIs can have significant social and economic impacts. Therefore, assessing the effectiveness of NPIs plays an important role in responding to the pandemic and other public health crisis.

During the pandemic, social media data and mobile network location data are the two main raw data sources to explore human mobility patterns according to requirement of the study. The methods for calculating human mobility are divided into three main types: frequency of people movement, distance of people movement, and the number of visitors for one space. The first two methods, typically the OD matrix and ROG, used to require a large amount of data to achieve better results, also may loss some details in spatial analysis. The third method uses active user number on social media, or the geo-tagged posts count as a representation of human mobility in customized areas and periods. This method does not require the starting and ending points of user movement. Compared to Google Community Mobility Report, the results may differ somewhat because common social media data only records where people stay, but not how long they stay there.

In addition, some gaps in research are identified. The existing human mobility studies applied in Canada during the COVID-19 pandemic are mainly using Google Mobility Reports and mobile network location data in a large spatial scale. Jiang et al.'s (2021) approach of using geo-tagged tweets to observe human mobility within New York City can be drawn on and applied to the GTA. This thesis builds on their approach to observe human mobility changes along the severeness of the pandemic over a longer epoch, as well as the impact of policy updates on human mobility until the end of 2020.

Chapter 3

Human Mobility Analysis based on Geotagged Tweet

This chapter introduces the GTA as the study area of this thesis in Section 3.1. Section 3.2 describes the data sources of geotagged tweets, land use data, and Google Community Mobility Report. Section 3.3 details the general framework of the proposed methodology, including the data pre-processing approaches for both geotagged tweets and land use data, and the proposed methods of monthly and weekly spatial analysis and daily temporal analysis, respectively. Section 3.4 summarizes this chapter.

3.1 Study Area

The Greater Toronto Area, commonly referred to as the GTA, is an economic zone in southern Ontario, Canada. It starts from Burlington on the western shore of Lake Ontario, through the City of Toronto, and extends to the Clarington on the northern shore of Lake Ontario. From the lake shore northwards to the Township of Georgina and Township of Brock areas along Lake Simcoe. It consists of the City of Toronto, and its four surrounding regional municipalities, Halton Region, Peel Region, York Region, and Durham Region, 25 cities in total, which covers an area of approximately 7,125 square kilometers. According to the 2021 Census data, the GTA region has more than 6.76 million population, which is the most densely populated region in Ontario.

One of the reasons of the dense population of the GTA is influx of labor, attracted by the large number of job opportunities available. The GTA is the largest regional economy in Canada, contributing approximately one-fifth of Canada's Gross Domestic Product (GDP) each year. The City of Toronto is the second largest financial centre in North America, after New York, but it is the fastest growing financial center in North America. All of Canada's five largest banks have their headquarters in downtown Toronto. The financial services sector is the largest industry in Toronto that contributing more GDP and jobs each year than the average of all other industries (The Conference Board of Canada, 2020). It is also the center of business in Canada, with many companies locating their Canadian headquarters in the GTA. Tourism is also one of the key industries supporting economic development in the GTA. Since 2013, visitor arrivals and total spending have continued to trend upwards. In 2018, Toronto's Census Metropolitan Area (CMA) hosted over 40 million visitors who spent over $10 billion Canadian dollars (Tourism Economics, n.d.).

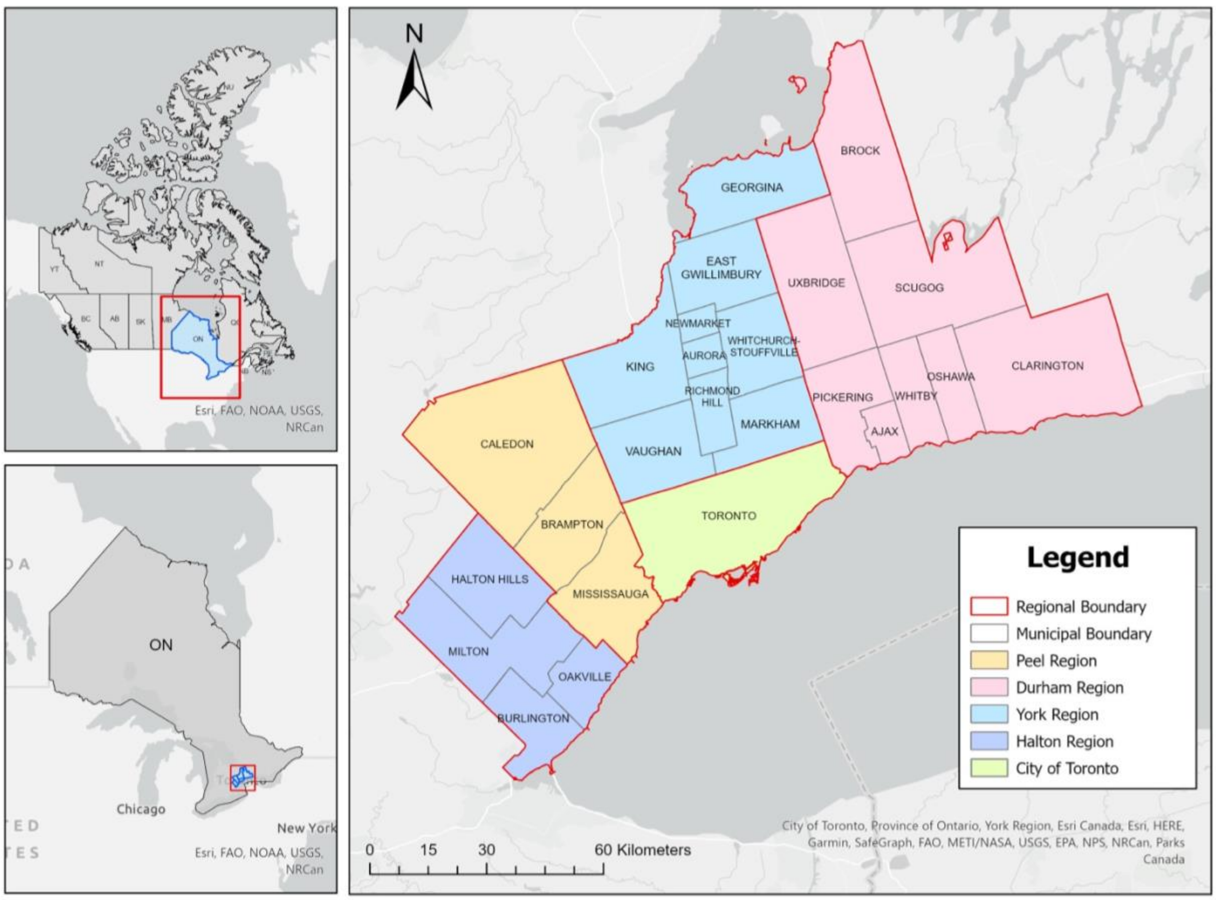


Figure 3.1 Map of the study area, the Greater Toronto Area in Ontario, Canada.

It is worth noting that the GTA area is defined as an economic zone and slightly different from the Toronto CMA. As a result, many official surveys do not provide complete coverage of the GTA. for example, Oshawa in the Durham Region is a separate CMA, but it is part of the GTA, while some areas of Simcoe County and Dufferin County are part of the Toronto CMA, but not the GTA. therefore, some GTA data can be roughly estimated through data from the Toronto CMA. To be more specific, the 2021 Census data provides a total population within the Toronto CMA that is approximately half a million less than the sum of the populations provided by each region within the GTA.

During a pandemic, the GTA shows its vulnerability to risk. According to the Ontario Epidemiology summary, in early April 2020, more than half of all new cases were in the GTA (Public Health Ontario, 2020), and by early June 2020, the proportion is more than two-thirds (Public Health Ontario, 2020), while the total population of the GTA was less than half of the total population of Ontario. Compared to the GTA's population as a proportion of Ontario's population, the number of new cases and population was disproportionate**.** In 2020 alone, more than 134,000 GTA residents have been infected by the COVID-19 pandemic, with more than 45% of confirmed cases reported from the City of Toronto, more than 30% from the Peel region, 13% from the York region, and the remaining 10% from the less populated Halton and Durham regions. The COVID-19 pandemic was directly or indirectly responsible for the deaths of over three thousand people. In addition, the social impact can take a long time to recover from. It is obvious that the COVID-19 pandemic causes a decrease in employment in the Toronto CMA, particularly in the retail and restaurant sectors. In the first half of 2020 alone, the unemployment rate in the City of Toronto almost tripled (Mowat & Raf, 2020). At the same time, the industrial structure of the GTA dictates that many people would take up remote work. The reliance on tourism and the number of student groups would lead to a reduction in tourist traffic.

3.2 Data

3.2.1 Geotagged Tweets Data

This study uses the Twitter API for academic research. The commonly used Twitter API randomly selects the 1% of all tweets that meet the filtering criteria. The Twitter API for academic research can access to a comprehensive geotagged tweet dataset. If Twitter users chooses to add geotags when they post the original tweet, the tweet is given a location ID and records accurate coordinate information if applicable. The tweets dataset was filtered by original tweets posted in 2020 and posted in 25 cities within the GTA. A total of 4,464,251 geo-tagged tweets are collected. However, those tweets require further cleaning because one user may post multiple times in one place. In order to protect user privacy, tweet information will not be displayed in the final results, but some user information such as user ID, will be retained during the data cleaning phase in order to clean up excess data. Whereas several temporal scales are designed in spatiotemporal analysis of human mobility changes based on land uses, for monthly analysis, data from January and February are used to calculate the baseline value. For the weekly and daily analysis, data for the five weeks from January 3 to February 6 will be used for baseline calculation. This helps to compare with the Google Community Mobility Report more conveniently. The rest of the data will be used for the calculation of in human mobility changes.

3.2.2 Land Use Data

Land use data for the GTA area is not aggregated at the regional level. Therefore, it needs to be obtained from the city's open data portal, the ArcGIS online portal, or other sources. The quality of the data may vary at different municipalities and requires some basic operations such as clip, merge, etc., before analysis. Most cities do not make a clear distinction between land use data and zoning data. Therefore, the term "zoning" is used as an interchangeable keyword when searching for land use data. For the purposes of this study, land use types will be categorized into residential, commercial, workplace, open space and transport based on the primary function of the land, with reference to the By-law Zoning guidelines provided by each city. Some cities have opened up free access to their land use data, but upon review these data may be too detailed, for example by dividing residential areas into more detailed parcels by density or main building type. In this study, land use types are classified according to the function of the land only, and those categories that are too detailed needed to be consolidated. Zoning data for most municipalities is collected through the above process, with one exception, Township of Brock does not have interactive zoning data available to the public yet, nor does it have access to integrated data from other common sources over a larger area. In addition, some of the protected areas within the city and its surrounding buffer zones are under the authority of other departments that have taken over from the municipality. Therefore, these areas are not included in the city's land use data. Table 3.1 shows the summary of original land use data into different regions.

Table 3.1 Summary of original land use data in Greater Toronto Area and 5 regions.

Region Name	Number of Polygons	Area (km^2)
City of Toronto	11,444	519.71
Peel Region	16,007	1,186.36
York Region	54,697	1,685.78
Halton Region	7,940	870.97
Durham Region	33,352	1,989.83
GTA	123,440	6,285.08

3.2.3 Google Community Mobility Report

Google publishs mobility reports covering 135 countries or regions worldwide, but most reports are limited to the spatial scale of the country or region (Google, 2022). Fortunately, data at the regional level for Canada is available through filtered data tables. The report is generated by comparing daily visitor number and baseline data for different location during the pandemic. The baseline data uses seven values corresponding to the day of the week instead of a fixed value. Each value is the median of each day of the week from January 3 to February 6, 2020, five weeks in total. This dynamic baseline data avoids the need for day-to-day comparisons, especially between weekday and weekend. In Google's Community Mobility Report, human mobility data are grouped into six categories of places, including retail and entertainment, grocery and drug shops, parks, transit stations, workplaces and residences. The report provides the percentage change in human mobility for each type, which does not mean that large changes correspond to more visitors and vice versa.

3.3 Proposed Methodology

Figures 3.2 and 3.3 show the flowchart to introduce the process of both spatial and temporal analysis of human mobility in different time units, respectively. Figure 3.2 presents the data preprocessing and analysis approaches of monthly and weekly spatial analysis. As shown in Figure 3.2, the input and output datasets of monthly spatial analysis are marked in deep blue, and those for weekly spatial analysis are in light blue. Figure 3.3 represents the process of daily temporal analysis. As shown in Figure 3.3, the green parallelogram means the input tweet datasets for both regional analysis and the GTA analysis. Light violet parallelogram shows the output of regional analysis, and the deep violet parallelogram shows the output of the GTA analysis. As shown in both Figures 3.2 and 3.3, the white polygons show the raw data and datasets that are preliminary organized. Red parallelogram means the input land use dataset for spatial join. Yellow rounded rectangles are the process of analysis.

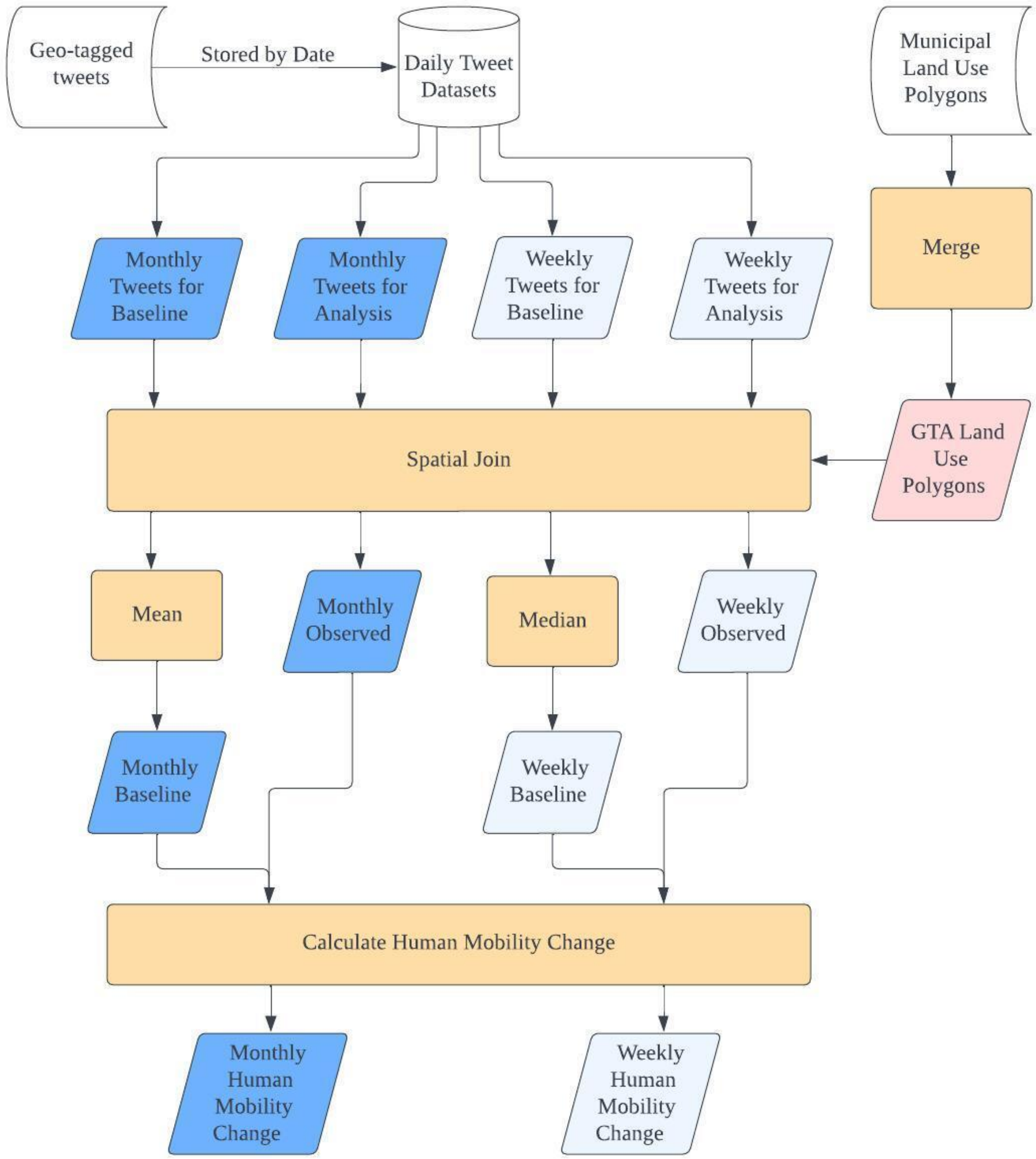


Figure 3.2 Flowchart of monthly and weekly spatial analysis based on land use polygons.

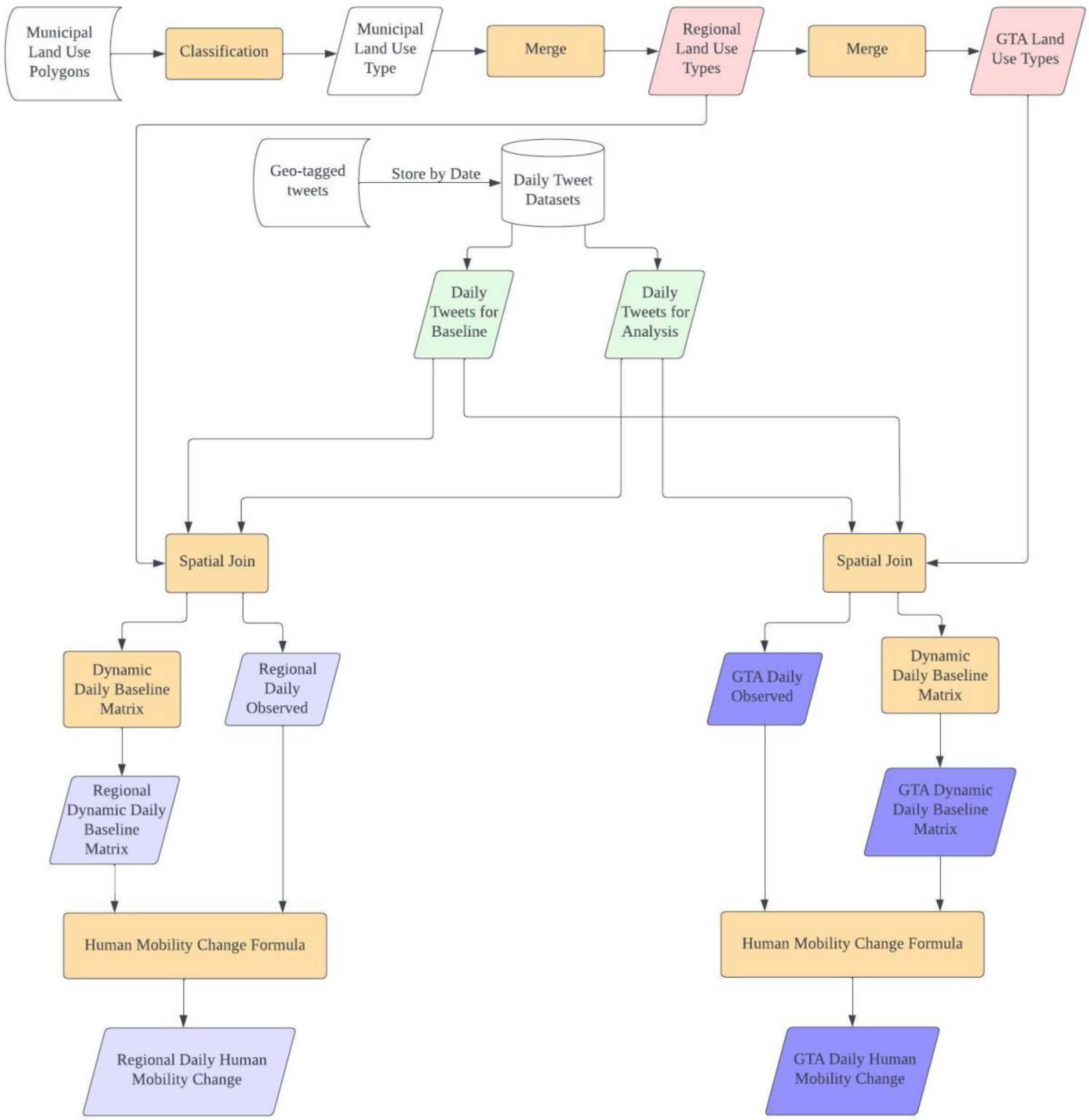


Figure 3.3 Flowchart of daily temporal analysis based on land use types.

3.3.1 Geotagged Tweets Data Processing

All geotagged tweets posted in the GTA in 2020 are collected using the Twitter Academic API, including many duplicate tweets and tweets posted by bots. They are stored in the respective datasets based on their posting dates. In order to eliminate the influence of these bots and commercial accounts on the results, collected data is required to be pre-processed before they used in analysis. There is only one tweet posted by a user in the same place will be remained in a daily geo-tagged tweet dataset. To be more specific, if a user posts multiple tweets from the same location in a single day, only one of them will be used in following analysis, which represents a user who was active at that location. If a user posts more than one tweets from multiple locations in one day, one tweet from each site will be held, which infer that the user had visited multiple locations during the day. This step uses the 'Remove Duplicates' tool in Microsoft Excel. After this step, a total of 1,463,790 tweets are remained. The next step is to transform the tweet location information into coordinates that can be easily mapped within the GTA's land use polygons. A part of the tweets is posted with accurate coordinates, which can map directly into the land use polygon. More tweets only have an encoded location ID with mixed number and alphabet, which need to find the appropriate JSON file provided by Twitter, then it can extract accurate coordinates of the location centroid from it. The location information contained in tweets may change over time. A small number of locations available when posting may become inaccessible when querying. For example, some shops and restaurants are replaced quickly are registered when the user posted the tweet with this location in 2020, but when querying the coordinates, the location information is no longer available. These tweets that cannot acquire the coordinates will also be deleted. Once this step has been completed, all tweets should have a pair of accurate coordinates. Next, tweets will be mapped to the land use polygon. A small percentage of data is deleted due to user error or system error. For example, confusing Richmond Hill in the York Region with Richmond in Vancouver. For those tweets that cannot be mapped to any polygon will be excluded from the spatiotemporal analysis.

3.3.2 Land Use Data Processing

For monthly and weekly spatial analysis of human mobility, land use polygons are used as a minimal spatial separation to project geotagged tweets and calculate human mobility changes. Therefore, the land use data collected for each municipality is simply merged in ArcGIS Pro except the lake polygons.

For daily temporal analysis, land use types for each region and the whole GTA are proposed. According to Section 3.2.2, some municipalities provide too detailed land use information to fit the requirement. Therefore, the municipal land use data need to be reclassified to primary land use types first. In detailed, Mississauga splits their residential areas into different building structures such as single-detached house, semi-detached house and so on. Milton classifies their residential zones by different residential densities from low to high. In daily temporal analysis, all those areas are reclassified as residential areas based on their primary purpose. After reclassified, there are several land use types, including residential areas, commercial areas, employment areas, industrial areas, institutional areas, open spaces, rural areas and so on. After reclassification, each city will have some or all of these classes. These reclassified land use types are then merged to each region and then to the entire GTA for counting geotagged tweets in each class. Figure 3.4 shows the distribution of main land use types in different regions and the GTA.

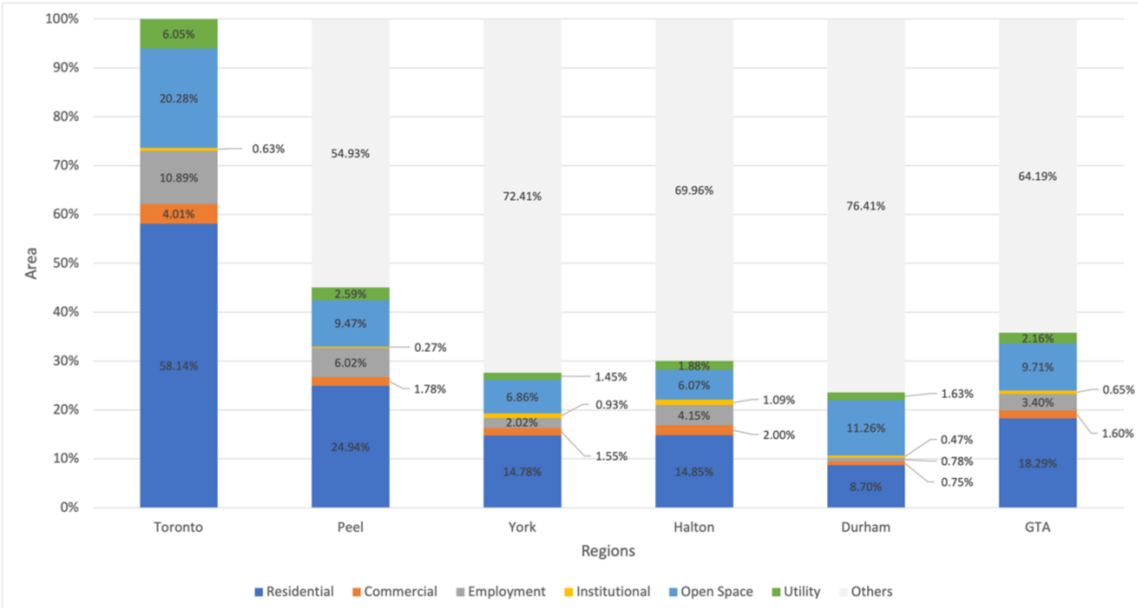


Figure 3.4 Land use distribution in the Greater Toronto Area.

3.3.3 Spatial Analysis of Human Mobility Changes

Monthly Analysis

The spatial analysis of human mobility is divided into two scales: monthly and weekly. The total number of users within each land use polygon for January and February need to be calculated independently first, then calculate their average that the sum of January tweet number and January tweet number divided by two. The total number of monthly tweets for the remainder of 2020 is used to calculate the change in human mobility. The total number of monthly tweets for rest of the month in 2020 are used to calculate the change in human mobility. The monthly tweets for the remaining months are spatial joined to the land use polygons. Then use the following formula to calculate the monthly human flow change for each polygon:

$Monthly\ Mobility\ Change =$

$$\begin{cases} \dfrac{monthly\ observed\ value - \ monthly\ baseline\ value}{monthly\ baseline\ value} \times 100\%, & if\ baseline\ \neq 0 \\ 0, & if\ baseline\ = 0\ and\ observed\ value = 0 \\ 100\%, & if\ baseline\ = 0\ and\ observed\ value\ \neq 0 \end{cases} \tag{3.1}$$

Weekly Analysis

Weekly baseline values are calculated using weekly total tweet number for five weeks between January 3 and February 6, 2020. Then spatial join the weekly tweets to the land use polygons and get join count for each polygon. The middle value from those five values for each polygon is the baseline value for weekly analysis. The weekly spatial analysis of human mobility covers seven stages, 17 weeks in total. Those time windows are chosen based on the policy released by Ontario government in response to the COVID-19 pandemic, which consider stages of the first outbreak in March 2020, recovery period between May and July, the second wave from September until mid-October, the implementation of colour-coded COVID-19 system in November, and finally the province-wide lockdown after boxing day (December 26). The specific timeline and key events are listed in Table 3.2. Overall, the policies in the Stage 1 are strict, people's production and life are shut down for two weeks. The policies from the Stage 2 to the Stage 4 are gradually relaxed. Stage 5 and 6 have moderately strict policies. The lockdown order from Stage 7 is strict again.

Table 3.2 17 Weekly time windows in 7 stages and important events.

Stage	Week	Date	Important Events
1 First outbreak	1	Mar 13 to 19	- March Break starting from Mar 16 - A state of emergency in Ontario declared on Mar 17 - Provincial parks shuttered on Mar 19
	2	Mar 20 to 26	- School continuously closed for two weeks - Non-essential business closure for two weeks
	3	Mar 27 to Apr 2	- School & non-essential business continuously closed - Shutdown of all outdoor recreation amenities
2 Initial Recovery	4	May 8 to 14	- Provincial parks opened on May 11 - Retail stores allowed curbside pickup on May 11
	5	May 15 to 21	- Campgrounds, marinas and golf courses opened on May 16 - More businesses reopened on May 19
	6	May 22 to 28	- Warmer weather caused people clustered in parks
3 Stage 2 of Recovery Plan	7	Jun 12 to 18	- Region of Durham, York & Halton moved to Stage 2
	8	Jun 19 to 25	- Peel Region & Toronto moved to Stage 2
4 Stage 3 of Recovery Plan	9	Jul 24 to 30	- Region of Durham, York & Halton moved to Stage 3
	10	Jul 31 to Aug 6	- Peel Region & Toronto moved to Stage 3
5 Second Outbreak	11	Sep 18 to 24	- The number of people allowed at private gathering limits in Toronto and Peel Region reduced on Sep 17, and expanded to whole Ontario on Sep 19
	12	Sep 25 to Oct 1	- Food and drink business asked to stop serving between 12-5am on Sep 25 - People gathered on Wasaga beach for a car rally on Sep 26
	13	Oct 2 to 8	- Restricted custom number in bar, restaurant, gym etc. on Oct 2
	14	Oct 9 to 15	- No indoor dining, closure of gyms and other recreation places starting from Oct 9
6 Colour-coded System	15	Nov 6 to 12	- Durham & Halton Region in yellow, Peel & York Region in Orange on Nov 7 - Toronto in red on Nov 11
	16	Nov 13 to 19	- Durham Region moves to orange; Toronto, Peel Region, York Region & Halton Region in red on Nov 13
7 Lockdown	17	Dec 25 to 31	- Christmas Holiday (Dec 25) - Provincial-wide Lockdown from Dec 26

The observed tweet number within each time window are spatial joined to the land using polygons, resulting in a weekly sum of tweets number on each polygon. The weekly change of human mobility in percentage is calculated from the following equations:

$$Weekly\ Mobility\ Change = \begin{cases} \dfrac{weekly\ observed\ value -\ weekly\ baseline\ value}{monthly\ baseline\ value} \times 100\%, & if\ baseline\ \neq 0 \\ 0, & if\ baseline\ = 0\ and\ observed\ value = 0 \\ 100\%, & if\ baseline\ = 0\ and\ observed\ value\ \neq 0 \end{cases} \quad (3.2)$$

It is worth nothing that when the baseline value is zero, but the observed value does not equal to zero, the equation does not make sense because the denominator is equal to zero. In this case, the change in percentage is assigned to 100%. For another case, if a polygon has a baseline value equal to zero, and the observed value also equal to zero. It is not difficult to understand that its human mobility has not changed during this period. Therefore, the change in percentage should be equal to zero. This idea will also be used in temporal analysis in below.

3.3.4 Temporal Changes in Human Mobility based on Land Use Types

The baseline data used in this study comes from the same five weeks from January 3 to February 6, 2020, as the weekly spatial analysis, but is calculated differently than in the previous subsection. This section uses the same matrix from the Google Community Mobility Report, with seven values corresponding to the dynamic baseline values for a particular land use type on a weekly cycle, resulting in a 6 × 7 matrix for each week, where 6 represents the six land use types of interest to this study and 7 represents the seven days of the week (Eq. 3.3).

$$Weekly\ Baseline\ Matrix_n (n \in [1,5]) = \begin{pmatrix} RM & RTu & RW & RTh & RF & RSa & RSu \\ CM & CTu & CW & CTh & CF & CSa & CSu \\ WM & WTu & WW & WTh & WF & WSa & WSu \\ IM & ITu & IW & ITh & IF & ISa & ISu \\ OM & OTu & OW & OTh & OF & OSa & OSu \\ UM & UTu & UW & UTh & UF & USa & USu \end{pmatrix} \tag{3.3}$$

where the first letter for each value in the matrix means the land use type: R – Residential Areas, C – Commercial Areas, W – Workplaces, I – Institutional Areas, O – Open Spaces, U – Utility and Transportation. The second letter to the third letter (if applicable) means the day of the week from Monday to Sunday, respectively. Since there are 5 weeks for baseline calculation, it produced 5 matrices in total. The final baseline matrix is a 6 × 7 matrix where each element in the matrix is the median of five weekly matrices.

Daily human mobility changes are calculated using the same theory with monthly changes and weekly changes, however, the baseline value for each day is not equal. Eq. 3.4 introduces two new variables, land use type and day of week, to reflect the dynamic baseline algorithm in human mobility temporal analysis. Another thing needs to be reminded is that the daily human mobility change analysis is based on the different land use types instead of land use polygon. Therefore, the sum of tweet number should be calculated based on each land use type before calculating the change in percentage.

$$Daily\ Mobility\ Change = f(x)$$
$$= \begin{cases} \dfrac{observed(LU_i D_j) - baseline(LU_j D_j)}{baseline(LU_i D_j)} \times 100\%, & if\ baseline\ value \neq 0 \\ 0, & if\ baseline\ value = 0\ and\ observed\ value = 0 \\ 100\%, & if\ baseline\ value = 0\ and\ observed\ value = 0 \end{cases} \tag{3.4}$$

The results of the temporal analysis of human mobility are presented in line charts with a trend line where the horizontal axis is the date, the vertical axis represents the percentage change in human mobility for each land use type.

The regional data obtained from the Google Community Mobility Report are used to create line charts based on the different regions to show trends in human mobility. The same or similar class are selected for comparison with the trends derived using the Twitter data.

3.4 Chapter Summary

In this chapter, the methodology of spatiotemporal analysis of human mobility change based on land use data is presented in detail. The workflow of spatiotemporal analysis includes three main steps: data preprocessing, spatial join, and human mobility change calculation. In data preprocessing, geotagged tweets are stored in different datasets based on their posting time. Land use data are reclassified and merged to fit the requirement of spatial and temporal analysis in ArcGIS Pro. After spatial join the geotagged tweets with land use polygons or types, the baseline value of monthly, weekly, and daily analysis is calculated using mean, median, and dynamic baseline matrix. Finally, the human mobility change is calculated using proposed equations.

Chapter 4

Results and Discussion

This chapter presents and discusses the results obtained using proposed methodologies. Section 4.1 presents the monthly spatial analysis results and discusses how the human mobility change among the pandemic severeness. Section 4.2 shows the weekly spatial analysis results in 7 stages with different policies. Furthermore, Section 4.3 presents the daily temporal analysis results. The results are compared to Google Community Mobility Report and discussed the possible reasons of differences in this section. Section 4.4 summarizes this chapter.

4.1 Monthly Spatial Analysis based on Land Use Polygons

The results of the monthly analysis of human mobility based on land use polygons in the GTA are shown in Figures 4.1 to 4.5. Each figure shows the results of two months from March to December 2020, respectively. A deeper blue colour means the human mobility decreased more, and a deeper red colour means the human mobility increased more.

According to the daily epidemiological summary published by Public Health Ontario (Public Health Ontario, 2020), the 2020 pandemic has three phases. The first outbreak peaks in April 2020. July and August are a recovery period with relatively few confirmed cases. The second outbreak begins in September 2020 and continuously growths until mid-January 2021 then gradually decreases. As shown in Figure 4.1, human mobility in most areas is basically remain the same during 2020, with the exception in Toronto, Vaughan and Brampton. The first response to the COVID-19 pandemic is introduced in mid-March 2020, when the human mobility in the GTA begin to fluctuate. Two residential areas in Vaughan and Richmond Hill experiences the change in human mobility exceeding 100%, otherwise, most urban areas except open spaces have their human mobility changes limited to 25% more or less. Rural areas and inner-city open spaces show increases in human mobility, in some places by more than 75% or even 100%. Toronto Pearson International Airport shows a 40% human mobility reduction in March. In late January 2020. the Canadian government issued a travel advisory, and Air Canada discontinued all direct flights to China.

According to the report, April is the most severe time of the first outbreak. Provincial parks, and outdoor recreation areas including beaches, playgrounds, sports fields, picnic areas and a portion of community parks are closed, but remain some other parks, trails, ravines and reserves for residents to walk. Human mobility in some conservation areas in the northern part of the GTA declines by more than 75%. In contrast, open spaces around the city have at least 50% increases in human mobility. In addition, downtown Toronto, Mississauga's downtown core, Oshawa's downtown, and the commercial areas of Vaughan and Richmond Hill have no less than 60% decreases in human mobility. No less than 25% decreases in human mobility are observed in the industrial areas of Vaughan and Richmond Hill. Reductions in human mobility are also observed in the work zones of Brampton, Mississauga and Oshawa, but the actual human mobility fluctuations may not be statistically significant due to the small baseline value. The human mobility at the transportation utility decreases very quickly during the first outbreak, almost doubling in April when human mobility decreases by almost 80% compared to March when it decreases by only about 40%. Even human mobility on provincial highways reduce to about 50%. This may be because people no longer need to commute to work or school and cancel their travel plans because of health concerns.

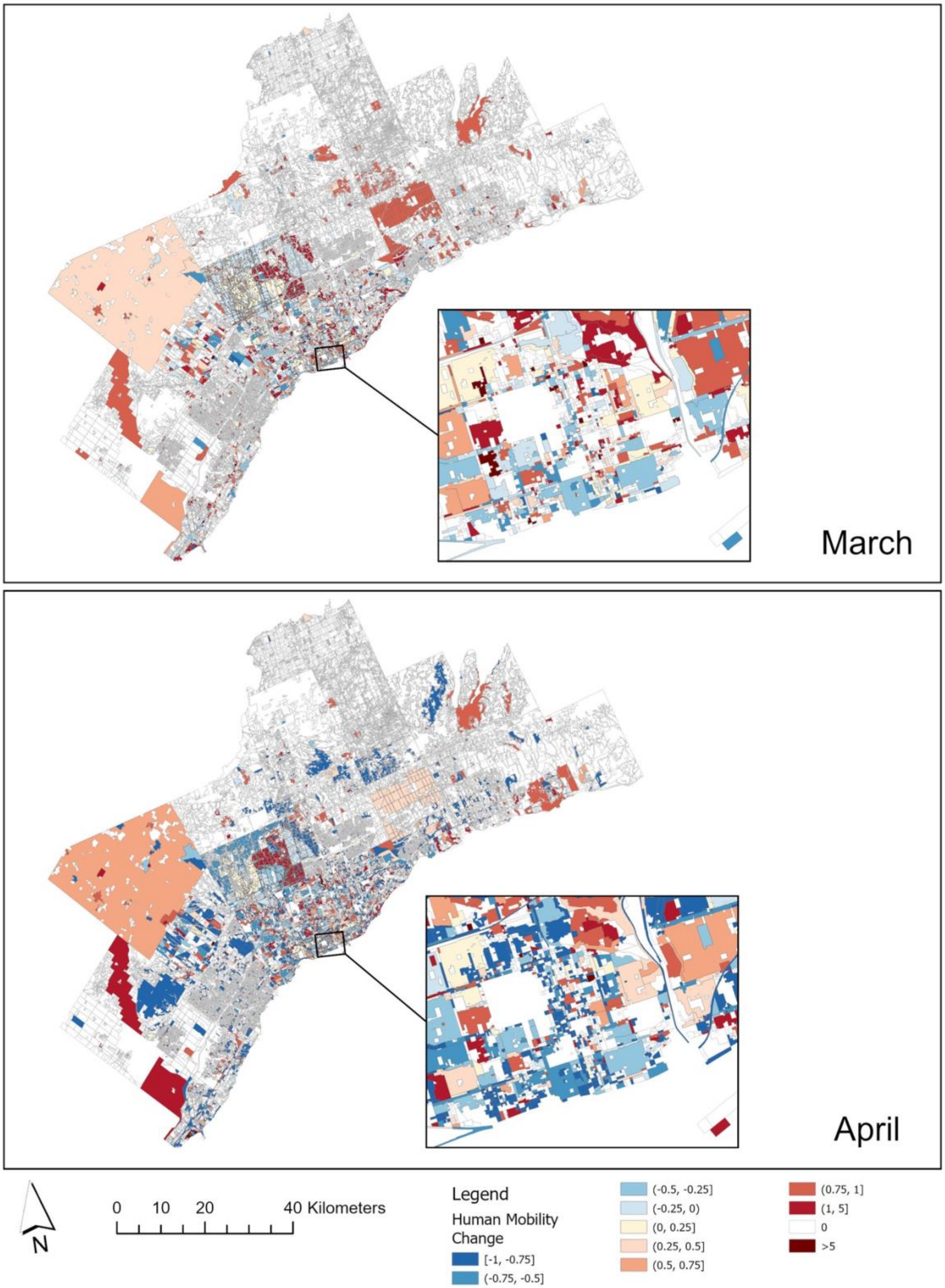


Figure 4.1 Monthly human mobility change pattern between March and April 2020.

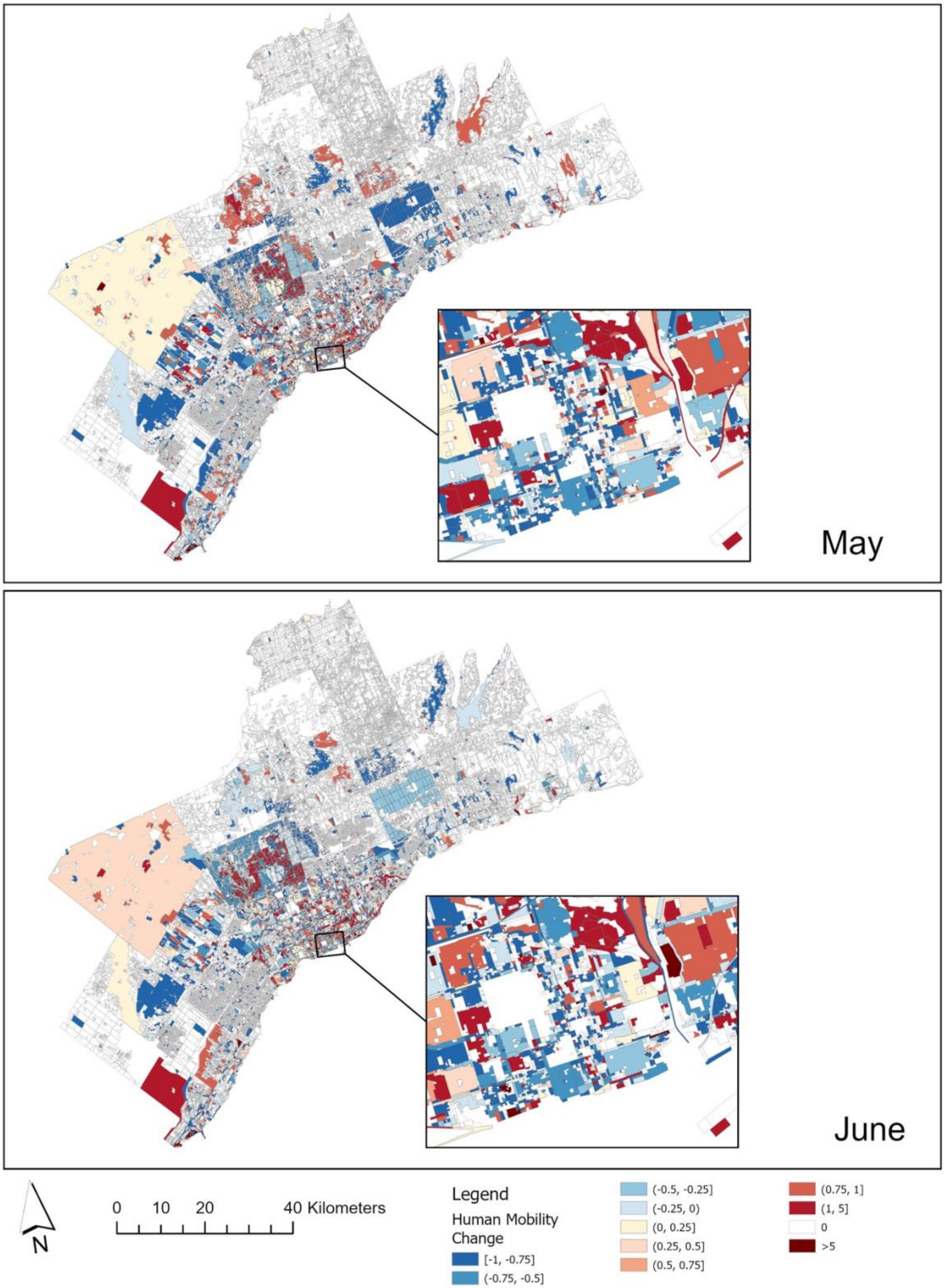


Figure 4.2 Monthly human mobility change pattern between May and June 2020.

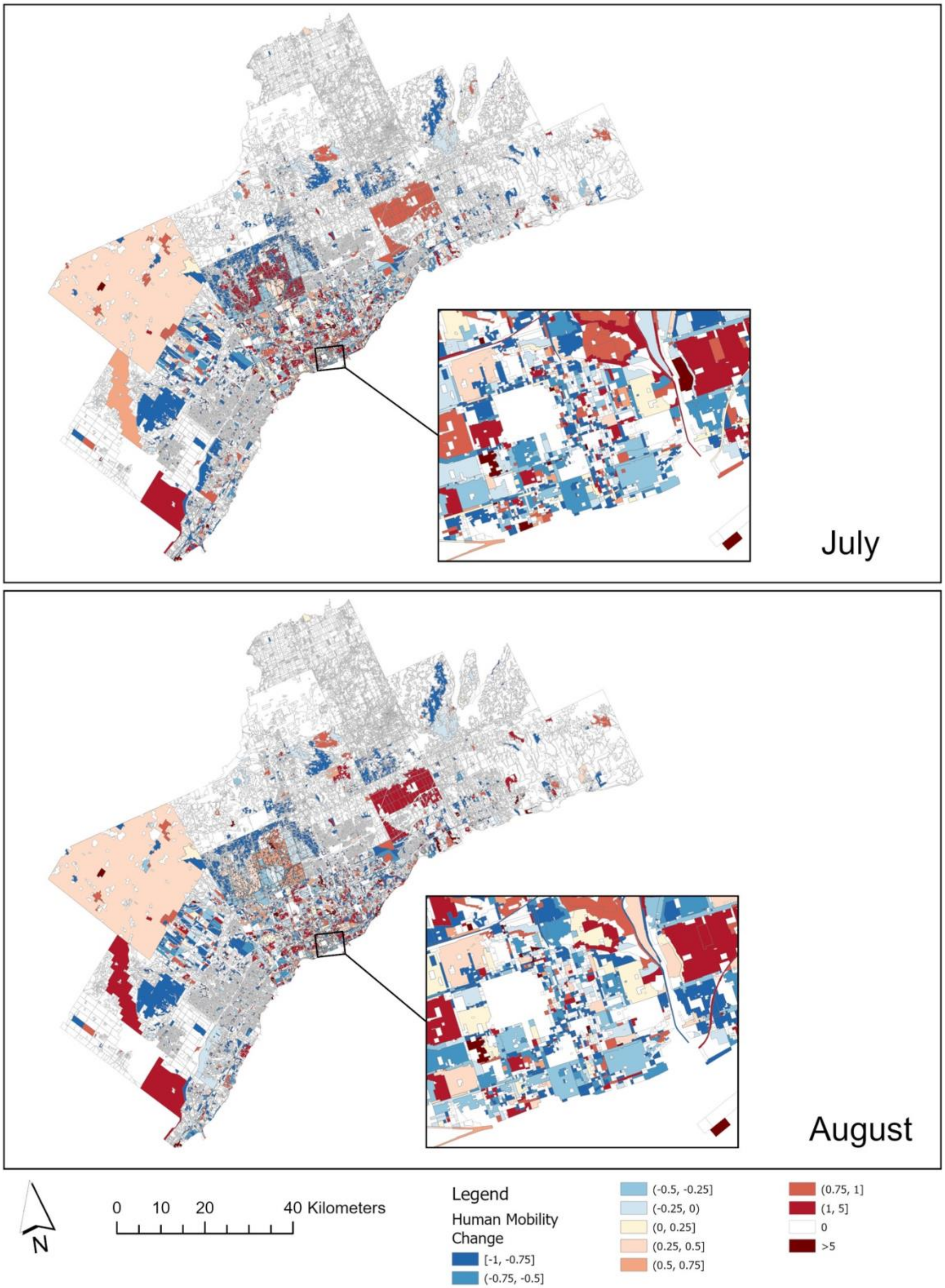


Figure 4.3 Monthly human mobility change pattern between July and August 2020.

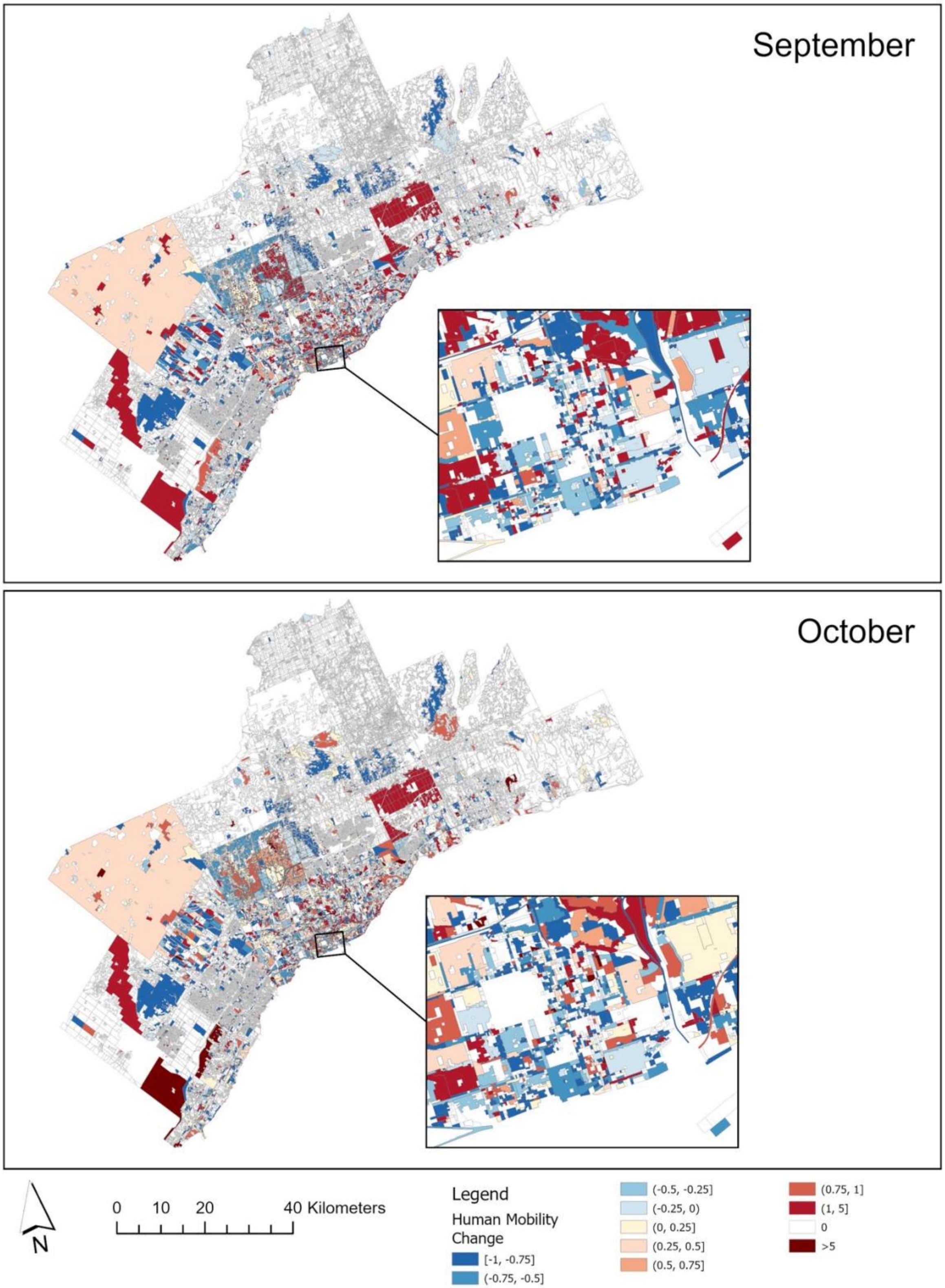


Figure 4.4 Monthly human mobility change pattern between September and October 2020.

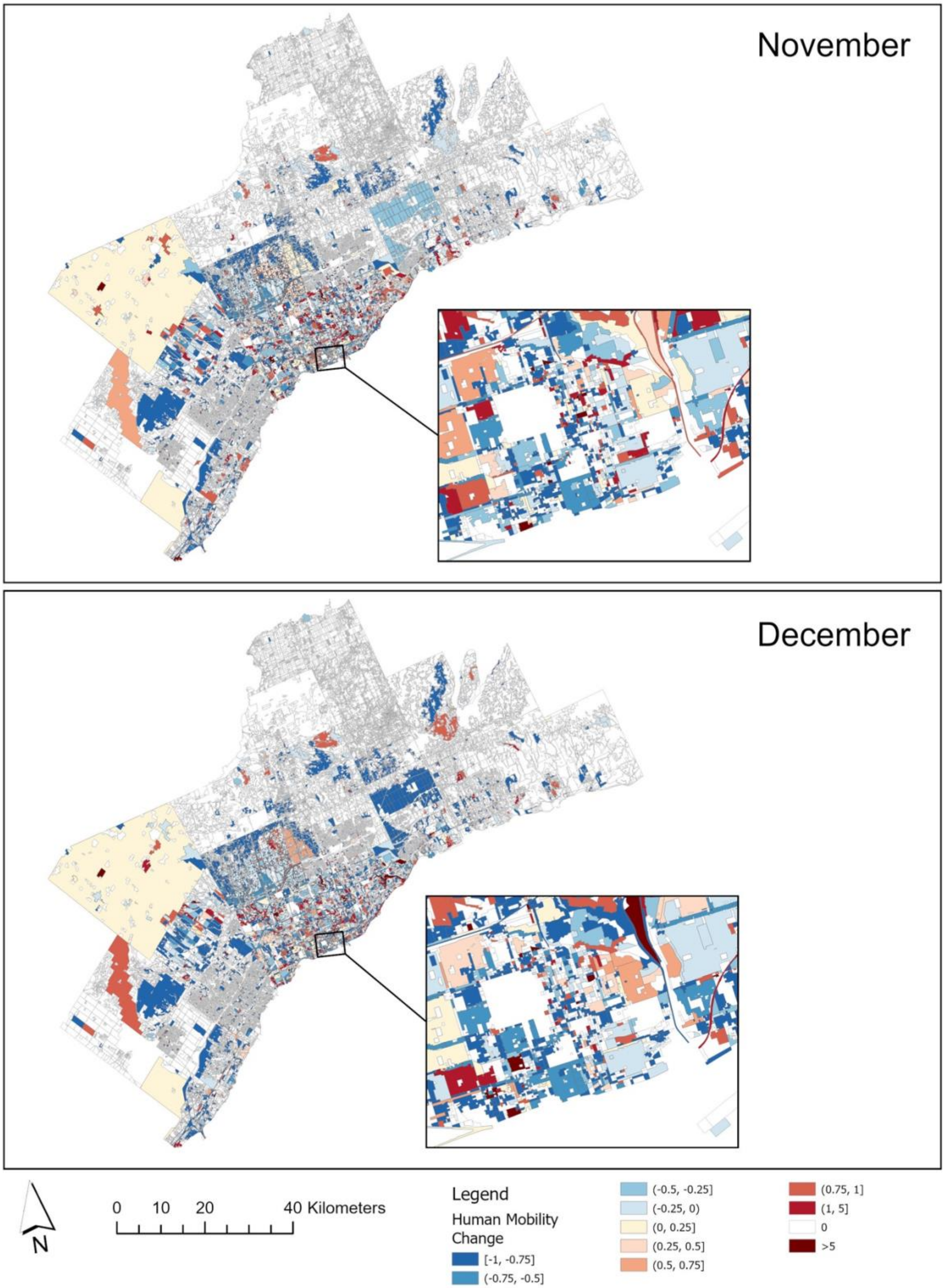


Figure 4.5 Monthly human mobility change pattern between November and December 2020.

The warmer weather is a shock to the virus in the summer. As a result, a recovery program starts in Ontario while the number of confirmed cases gradually declined. The recovery of human mobility in commercial areas receives some success in early July. It is observed that in downtown Mississauga, human mobility increases by about 20% in July compared to its April level, and by September is essentially at the same level as it before the outbreak. Human mobility in Vaughan's commercial areas reaches a stage peak in July but is still 43% lower than that before the outbreak. However, it does not continue to vary with the confirmed case number in the second half of the year, instead experiences a fluctuation and finally stabilized at around -56%. Human mobility in the Richmond Hill commercial area is more sensitive to the severity of the COVID-19 pandemic. After the April peak, human mobility recovers significantly in May, but shrink at the same time as a small rebound in June. As the number of confirmed cases fall in July and August, human mobility in the commercial area is slowly recovered. The lowest point of this period reaches in September, when roughly 40% of human mobility is restored. The reduction of human mobility in the downtown Toronto commercial areas also decreases gradually with the recovery plan, which can be seen to cross a color class over several months. Due to the scattered land use classification, it is not consolidated to calculate the trend.

As the weather warmed up, parks and green spaces become an ideally place for people to relieve the stress of the COVID-19 pandemic while keeping a social distance. In July, some environmental areas in the City of Toronto have a change in human mobility exceeding 500%. Another interesting observation is that human mobility on some agricultural lands increases from June to October and starts to fall back until November. The trend curve for the COVID-19 pandemic, on the other hand, has rising rapidly since September and has already surpasses the peak of the first outbreak in October. There are even some agricultural lands that had a growing human mobility during the first outbreak in April 2020. This shows that the difference between human mobility change patterns on agricultural land and the COVID-19 pandemic spread patterns cannot be explained simply by a potential lag time. After reviewing Ontario's harvest schedule (Foodland Ontario, n.d.), we find that June to October is the main harvest period for crops in Ontario. During this period, tourism and jobs developed from harvesting may be an important factor that determines the human mobility change patterns on agricultural lands.

The human mobility change patterns observed in December are broadly the same as it in April. However, without the restrictions placed on provincial parks and others in April, more environmental areas are observed to have increased human mobility in December. The majority of open spaces within the City of Toronto exhibits an increase in human mobility in excess of 100%. Unlike the performance of inner-city open space, agricultural lands experience a sharp decline in human mobility due to the harvest schedule. Also, although the second outbreak is much more severe than the first, the figure shows that the human mobility change in the Vaughan residential area gradually decreases from September onwards, whereas the human mobility change in the Richmond Hill residential area, just down the street, continue to increase from September onwards.

4.2 Weekly Spatial Analysis based on Land Use Polygons

Prior to this, we assume the human mobility change is not only be influenced by the severity of the COVID-19 pandemic, but also policies updates. Analysis on a monthly scale does not provide a strong explanation of changes in human mobility before and after policy implementation, and therefore a finer temporal granularity is needed to help understand the effectiveness of the policy. Figures 4.6 to 4.10 represent weekly human mobility changes over a seven observation windows based on the policy releases, for a total of 17 discrete weeks, between

March 13 and December 31.

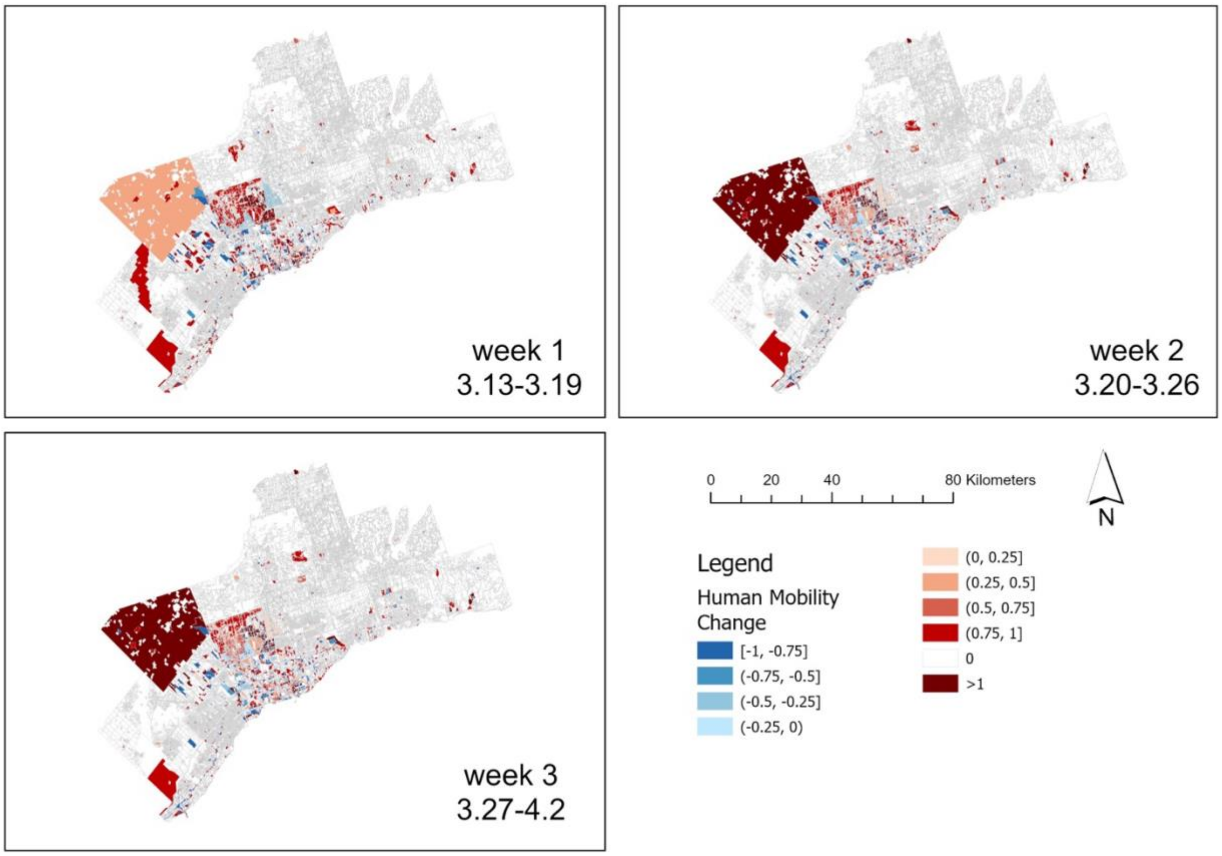


Figure 4.6 Human mobility changes in the second stage from March 13 to April 2.

Figure 4.6 presents the change in human mobility over the first three weeks. The first week shows an increase in human mobility in the open spaces and residential areas, mostly above 75%. There is a decrease of approximately 20% in human mobility in the commercial areas prior to the pandemic, with some exceptions, which may be due to the rush for necessities. Human mobility in the employment areas is also reduced from pre-pandemic period, but less significantly. The order to close provincial parks is issued at the end of the first week. It makes the human mobility in some open spaces decreases compared to the previous week, while others increase in the second week. Also, schools are continuously closed, and some non-essential businesses are shut down, forcing people to stay at home to stop the spread of the COVID-19 pandemic. Human mobility in institutional and commercial areas continuously declines to less than a half before the outbreak. More outdoor recreation facilities are closed during the third week. Human mobility at a beach in Scarborough is decreased, meanwhile, there is an increase in nearby residential areas.

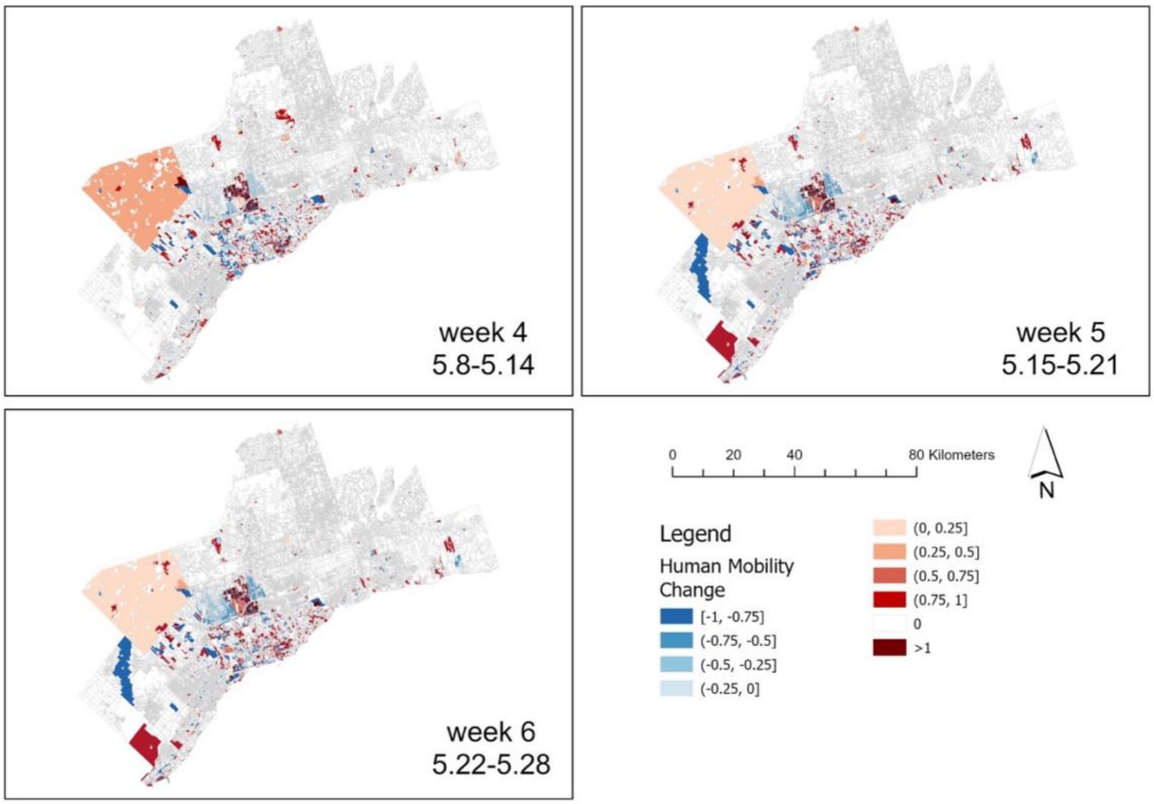


Figure 4.7 Human mobility changes in the second stage from May 8 to 28.

The fourth observation week allows retail stores to curbside pick-up, but human mobility in commercial areas is still far worse than it before the pandemic. Provincial parks are allowed to reopen, as a result, some open spaces receive a significant increase in human mobility, even surpassing the pre-pandemic level more than one times. More outdoor amenities are opened in Week 5, and beaches, parks with campgrounds become popular destinations, and even the highways that connect them have a significant increase in human mobility. In addition, human mobility in commercial areas gradually returns to around 40% of pre-pandemic levels during the week. Increased human mobility is in more residential areas within the City of Toronto in Week 6, but it is difficult to interpret this based on the policy of the recovery period. Thus, this anomalous increase may related to the discussions of the deaths of Regis Korchinski-Paquet (Special Investigations Unit, 2020) and George Floyd (The New York Times, 2022).

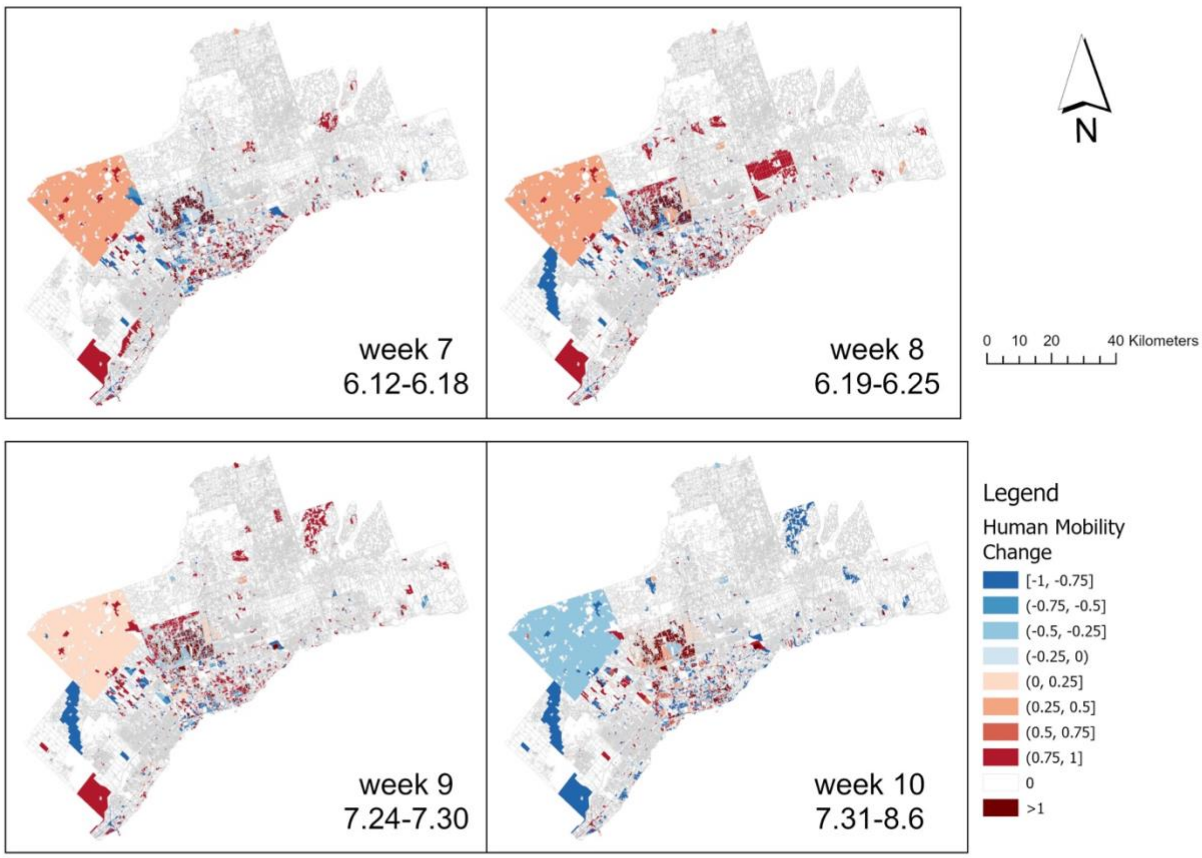


Figure 4.8 Human mobility changes in the third stage from June 12 to 25 (upper two) and it in the 4th stage from July 24 to August 6 (lower two).

The GTA is fully in Stage 2 in Weeks 7 and 8 and in Stage 3 in Weeks 9 and 10. Greater human mobility than pre-pandemic level is observed in the commercial area in some cities in the Durham Region, York Region and Halton Region in Week 7. The commercial areas of Toronto and Markham, Richmond Hill are also gradually recovering, but they are not reaching pre-pandemic levels yet. A trend towards increased human mobility in the commercial areas within Toronto is observed in Week 8. The highways through Toronto also recover to half of their pre-pandemic level of human mobility. Human mobility in some residential areas is slightly less than before the pandemic. By Week 9, the highways almost recover to pre-pandemic human mobility, and human mobility in the commercial areas of Vaughan and Richmond Hill have returned to nearly 60%. The highways through the City of Toronto have less human mobility in Week 10, but the highways to downtown Toronto show more human mobility. Downtown Toronto has more human mobility than in Week 9.

Human mobility increases in residential areas and open spaces during Week 11. Human mobility in commercial areas drops to 50% of its original level in Week 12. Highway 401 experiences an increase in mobility, but the roads and railways leading to downtown Toronto are emptier than before. Other cities also show a significant increase in residential area and a decrease in commercial area. Nevertheless, there are some rural areas in the Durham Region that received more visitors. The industrial areas of Vaughan experience a significant mobility decrease in Week 13. Residential areas in the City of Toronto and Brampton show an increase in mobility. However, the open spaces in western Toronto have decreases in human mobility. Week 14 features a somewhat different change. Human mobility decreases in the residential areas of Vaughan and the southern part of Toronto. However, it increases in the rural and environmental areas of the surrounding cities, as well as in open spaces within the City of Toronto.

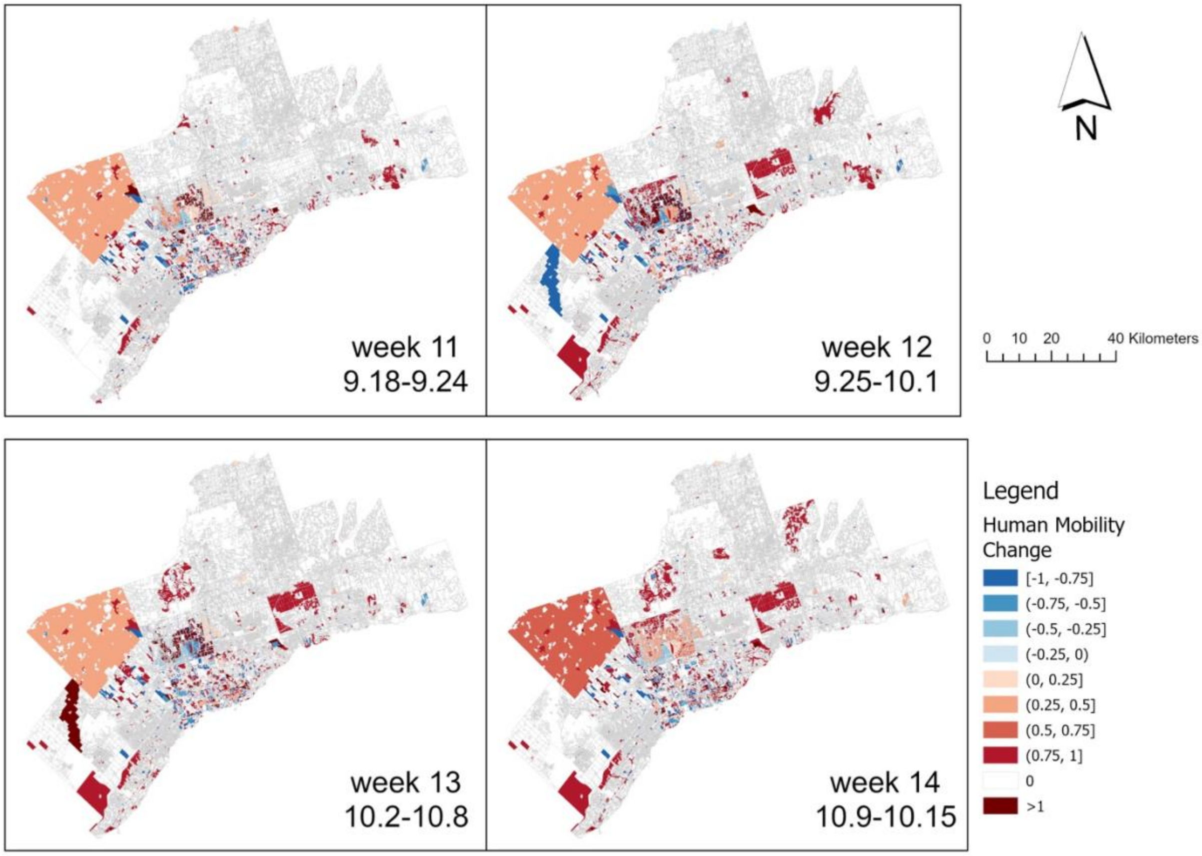


Figure 4.9 Human mobility changes in the 4th stage from September 18 to October 15.

There is still a human mobility growth in Week 15 in rural areas of the cities surrounding Toronto and in Toronto's open spaces. Human mobility in commercial areas declines to less than 50% of pre-outbreak levels. Less than 25% mobility is observed on main transportation routes in York Region and over 50% reduction in highway mobility within the City of Toronto, but over 25% growth in mobility is observed on Queen Elizabeth Way (QEW) within the City of Oakville. York Region and Peel Region residences show increases in mobility. Week 16 experiences a 50% decrease in mobility in the York Region's commercial areas. Human mobility in rural and environmental areas reduce compared to Week 15. More residential areas within the City of Toronto show an increase in human mobility. The QEW in Oakville experiences continuously increased human mobility. Some residential areas in Brampton no longer show an increasing trend in human mobility.

Week 17 is another independent observation period. Human mobility increases in some rural areas. Residential areas in Vaughan show an increase in human mobility over 100%. Commercial areas in the Durham Region show a decreased human mobility for more than 50%. More residential areas in the City of Toronto show a decrease in mobility than those showing an increase in mobility. Residential areas in Brampton show a decline in mobility as well.

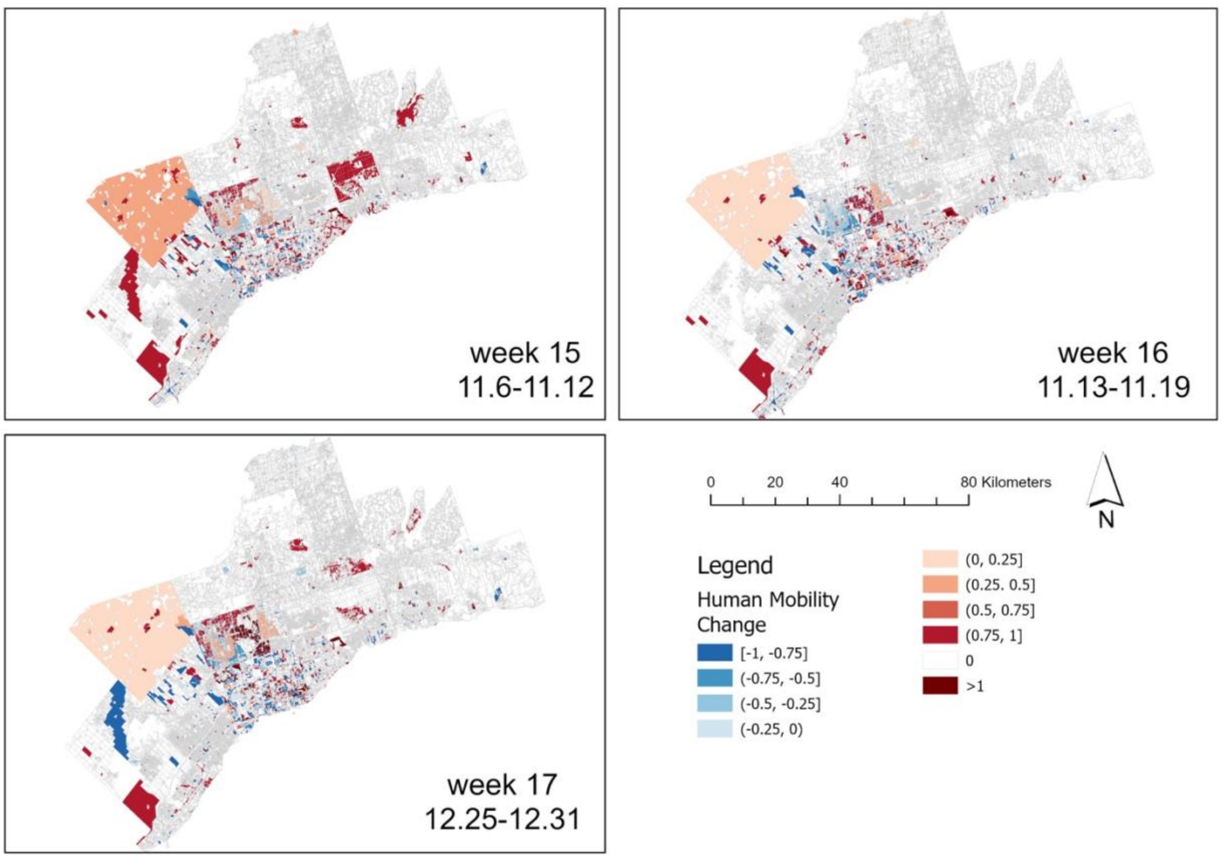


Figure 4.10 Human mobility changes in the 6th stage from November 6 to 12 (upper two) and it in the 7th stage from December 25 to 31 (lower one).

4.3 Daily Temporal Analysis based on Land Use Types

The daily human mobility change analysis uses a dynamic baseline matrix applied to the 6 major land use types in the GTA. Figure 4.11 shows the changes in human mobility for these land use types. The black line is the day-by-day human mobility derived using the Twitter data, and the red line is the trend line after Gaussian kernel smoothing, which is a type of weighted moving average that gives the values closer to kernel a higher weight when calculating the average. All trend charts in this section are presented in this style and will not be explained repetitively in the subsequent sub-sections.

Human mobility in residential areas rises 20% in the first outbreak, then declines and stabilizes around baseline levels from May to June. There is a slight decline of less than 10% on average from August to the November 2020, although there are fluctuations in November. There is a significant drop in December and there is no significant rebound until the lockdown after the boxing day. Human mobility in the commercial areas trends upward prior to the first outbreak but declined significantly during the first outbreak. After that, it remains at a slightly lower level than the baseline value, with an average decline of less than 10%. Both October and December see varying degrees of decline and bottlenecks until reaching the lowest point at Christmas. Human mobility in employment areas is reduced to less than 70% of the baseline value during the first outbreak. It then rebounds marginally as the number of new confirmed cases decrease but is still nearly 30% below the baseline value. With the growth of the second outbreak, human mobility in the employment areas decreases again until the end of 2020. Human mobility in the institutional areas declines significantly in late March, continuously declining to below 70% of the baseline value by the end of 2020. During the first outbreak, a significant decrease in human mobility in open spaces is observed, afterwards, a significant weather and severity driven change is observed, even weather becoming the primary driver instead of the pandemic severeness. This is similar to the trend observed in previous monthly analyses, where the warmer the weather, the more people engage in outdoor activities, and as temperatures cooled, human mobility in open spaces declines.

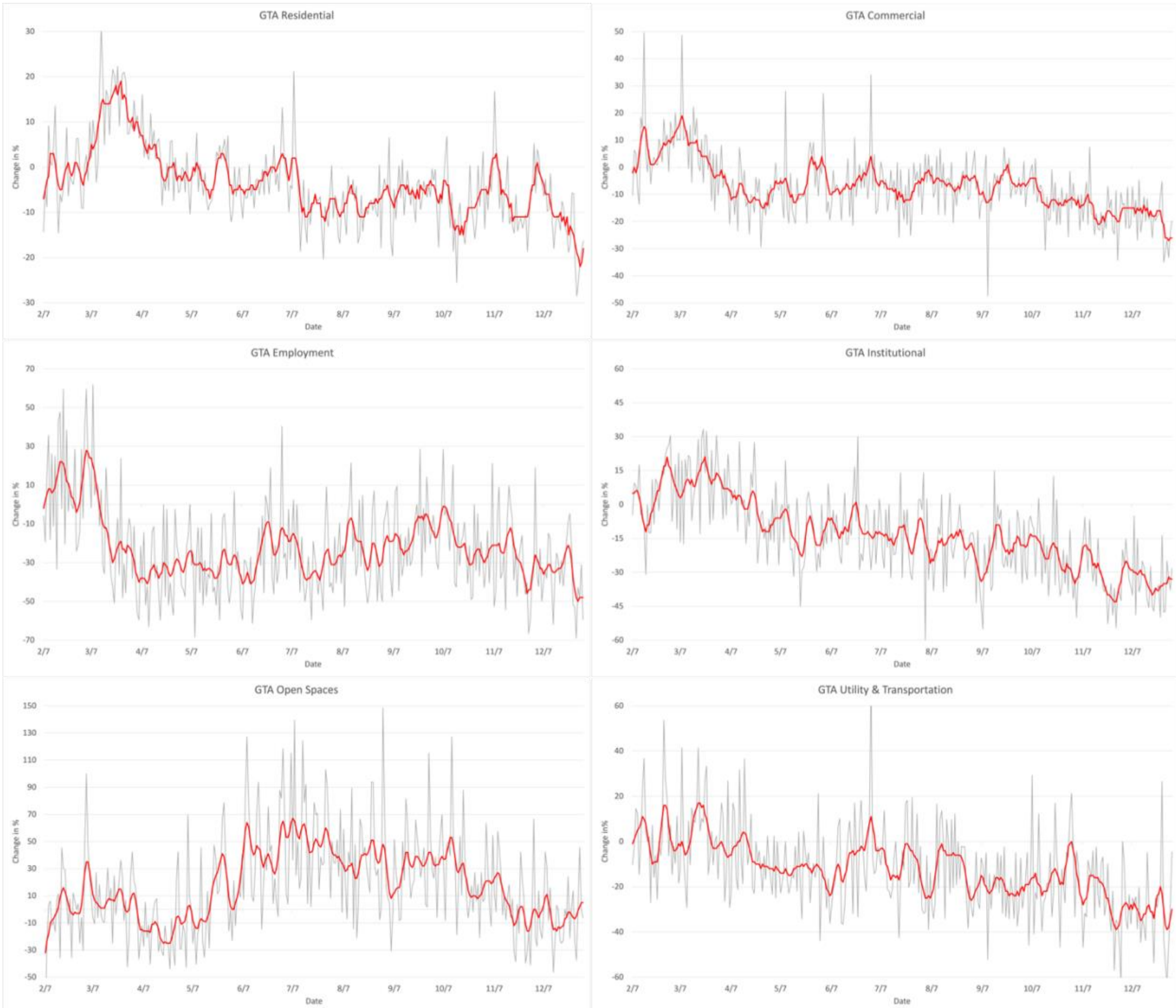


Figure 4.11 Daily Human mobility change in six land use types in the Greater Toronto Area from March to December 2020.

The changes in human mobility in the utility and transportation areas combine the trends in the commercial, employment, institutional, and open space areas. It shows an increase in human mobility in the second half of March and then decreases until May, which is consistent with the trend in institutional areas. There is a brief rebound from June to August, similar to the human mobility change in open spaces. Around 80% of baseline values are stabilized in September and

October, with a significant pickup from mid to late October. A small peak is observed at Thanksgiving and a large peak is observed at Halloween, presumably resulting in a different travel demand due to different holiday habits. December is a new low-mobility period, with a decrease of around 30% in human mobility in the utility and transportation areas. The increase is more likely to the pre-Christmas trend in the employment areas. The last week of 2020 shows a rebound in human mobility in utility and transportation areas, probably because open spaces become better choices of trip destinations after the lockdown order between December 26, 2020, and January 23, 2021.

4.3.1 City of Toronto

A total of four land use types in the City of Toronto, residential, employment, open space, and utilities and transportation areas, are selected to compare with the Google Community Mobility Report. Figure 4.12 illustrates the daily human mobility changes in Toronto from March to December 2020 derived by Twitter compared to the daily mobility changes provided by Google. The left column charts represent the human mobility change by land use type calculated from tweet data, and the right column charts represent the change provided from Google mobility data.

As shown in Figure 4.12, Twitter-derived human mobility in the residential areas trends upward at the beginning of the first outbreak but declines in April and then stabilizes consistently at about 120% of the baseline level for several months until November. There is a brief rebound in the week of the lockdown. A peak occurs in late-June, but it is not supported by Google's mobility report. It can be determined that this peak may be due to some hot topics that caused people to discuss more on social medias, which can only represent more stay-at-home people posting tweets during this period, but do not represent more people visited residential areas. In contrast, Google Community Mobility Report indicates that even though a downward trend began in April, the residential human mobility still 20% above the baseline value. A significant weather-driven change in human mobility can be seen in the rest of the year, reaching a valley in July and August and rebounding by December to slightly below the baseline level.

Human mobility in employment areas calculated from tweet data observes the same trends as in the Google report, only some different in ranges. Human mobility calculated using the number of tweets decreases by 40% in the first outbreak, while Google shows a 60% decrease in that period. Although human mobility tends to increase or decrease depends on the confirmed cases number after the first outbreak, human mobility derived by Twitter decreases on average by 30%, while human mobility provided by Google decreases by about 50% compared to the baseline value. Similar trends between Twitter-derived pattern and Google provided pattern are also found within the open space in Toronto, but the range of variation calculated using the tweet data is still smaller than that provided by Google.

Similar decreasing trends in human mobility are observed in the utility and transportation zones during the first outbreak, with less apparent weather-driven trends observed in the rest of the time. The uncertainty of the tweets data causes large fluctuations in the human mobility change patterns.

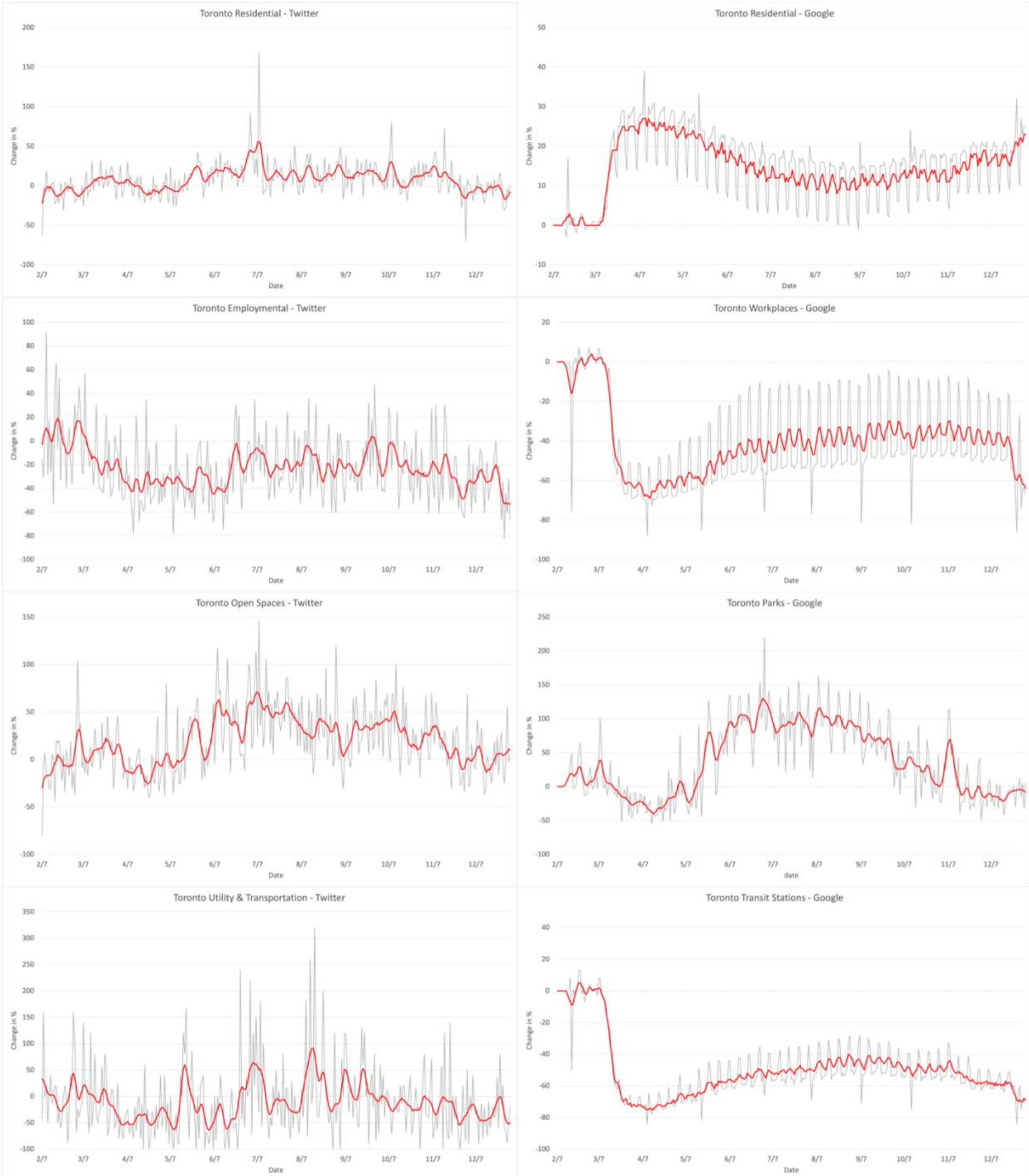


Figure 4.12 Daily human mobility change in four land use types in the City of Toronto.

4.3.2 Peel Region

The Peel Region also has four land use types, residential areas, employment areas, open spaces, and utilities and transportation areas, can be used to validate against the Google Community Mobility Report. As shown in Figure 4.13, similar trends of most land use types used for comparison are observed. The difference is that their range of variation is generally smaller than the data provided by Google.

A significant upward trend in residential human mobility in the Peel Region is observed during the first outbreak, but afterwards it declines continuously until 10% less than the baseline level. As the winter come, there is a slight increase in human mobility in the residential area, but it falls sharply again in December. Google mobility trend also shows a seasonal variation, it is always about 10% higher than the baseline value even in summer. The potential reason of this phenomenon is that people to lose something interesting in their lives, so the desire to share decreases tremendously, which leads to a lower human mobility in residential areas during the pandemic rather than the pre-pandemic level. Changes in human mobility in employment areas are closer to that provided by Google. Both observe a cliff-like decline during the first outbreak, after that a slight seasonal variation was found, but basically fluctuated around 50% of the baseline value. Human mobility in open spaces shows a significant downward trend during the first outbreak, also followed by a seasonal variation. Human mobility in open space calculated by tweet number is essentially equal to the baseline value from June to October, but the seasonal variation found by Google brings human mobility up to nearly twice the baseline value. Human mobility in the utilities and transportation areas experiences a significant decline during the first outbreak, then remains stable at less than 50% of the baseline value.

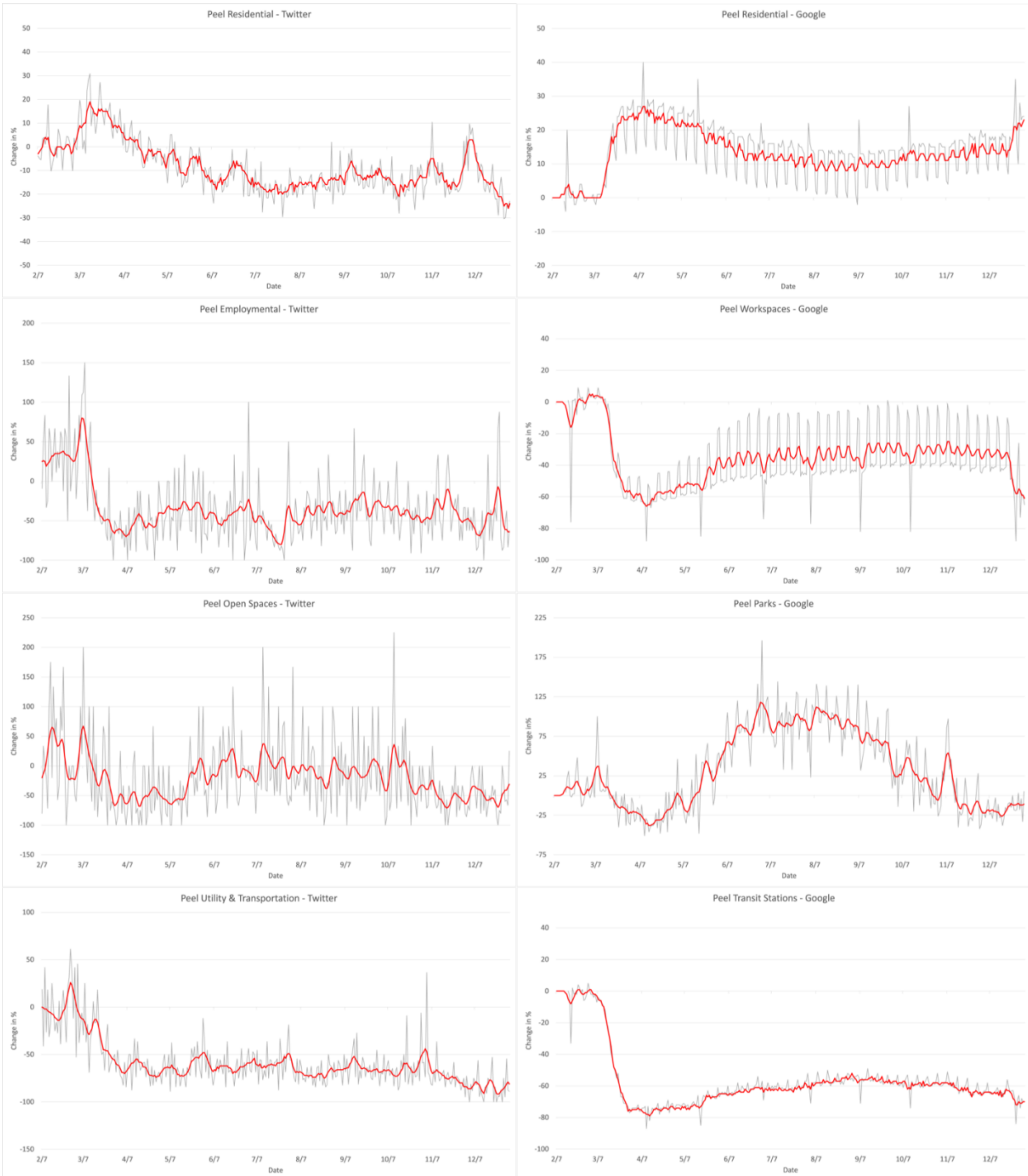


Figure 4.13 Daily human mobility change in four land use types in Peel Region.

The human mobility in the utility and transportation areas is related to travel demands, thus, the intensity of seasonal impacts on it is related to the seasonal variance in employment areas and open spaces. The pattern of human mobility change calculated using the number of tweets does not reflect strong seasonal variation in both employment areas and open spaces, so it is difficult to see the seasonal variation in utilities and transportation areas. In contrast, Google exhibits strong seasonal variances in the mobility change in employment areas and open spaces, the seasonal effects on utility and transportation areas can be found more easily.

4.3.3 York Region

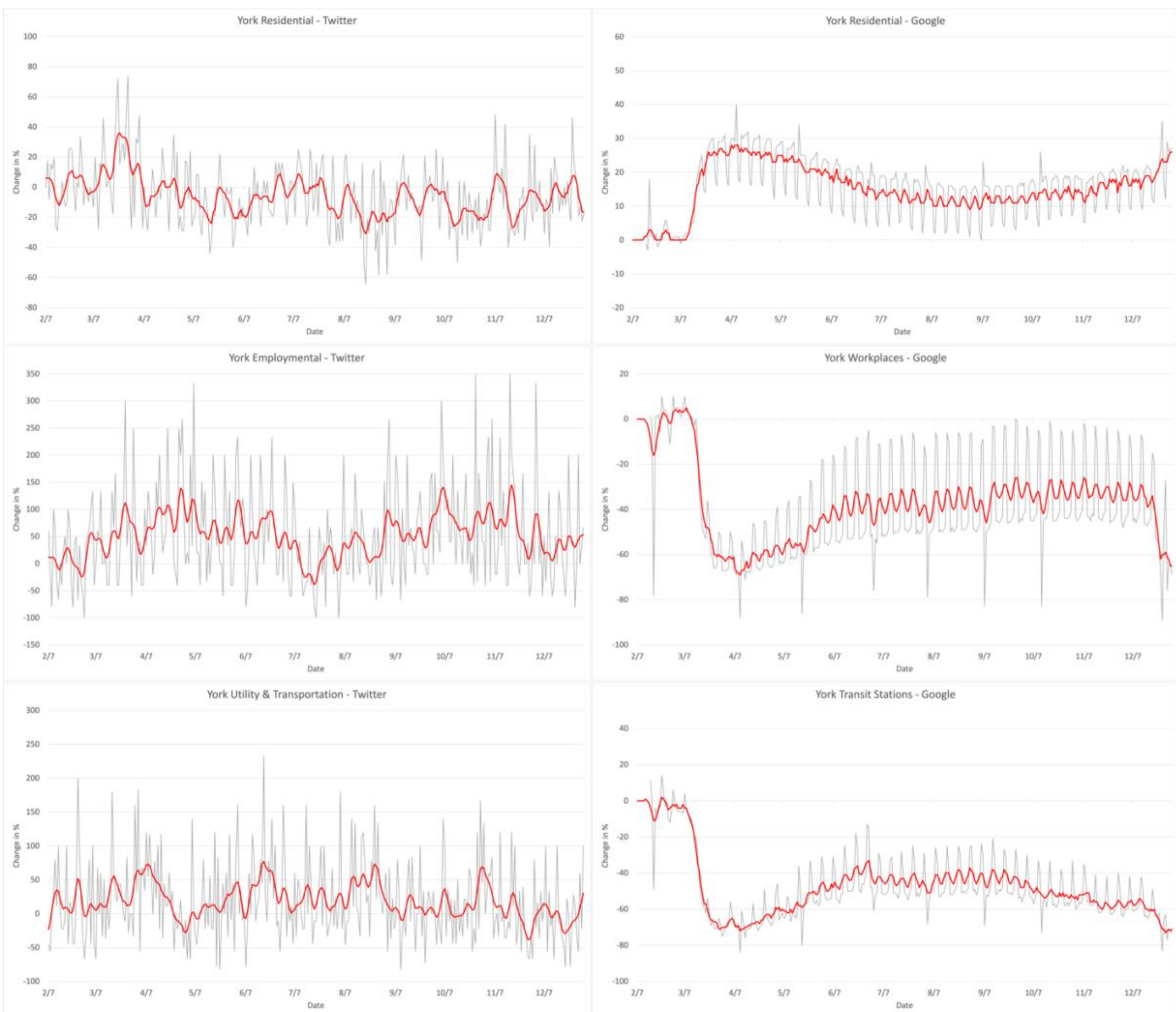


Figure 4.14 Daily human mobility change in three land use types in York Region.

There are three land use types in York Region that can be used for comparison with the Google Community Mobility Report, residential areas, employment areas, and utility and transportation areas, respectively. The changing patterns of human mobility in York Region and the comparison results are shown in Figure 4.14.

Human mobility in the York residential areas rises rapidly during the first outbreak, but then falls rapidly and remained below baseline values. No significant seasonal variation is observed. Two peaks in human mobility are found in the employment area, occurring in April and October, and a valley in July. In contrast, Google finds a more easily explained change pattern. There is a significant decrease during the first outbreak and then a gradual recovery to about 70% of the baseline value, with the Christmas holidays and lockdown causing a significant decrease in human mobility in the employment areas. Compared to other regions where have explainable human mobility changes in employment areas, this phenomenon may be due to the separation of industrial and employment areas in land use classification. Employment areas in other regions are typically larger than industrial areas, however, the industrial area in York Region is significantly larger than employment area. Therefore, this pattern may not be representative of overall working conditions of York Region residents.

4.3.4 Halton Region

A total of four land use types are selected for validation in the Halton Region, residential areas, employment areas, open spaces, and utilities and transportation areas. Figure 4.15 shows the variation in human mobility on those four land use types, with similar change patterns found on most of the land use type.

As shown in Figure 4.15, human mobility in the Halton residential areas rises 40% during the first outbreak, which is higher than the results from the Google Community Mobility Report. However, it does not remain at a higher level after the first outbreak, but slowly declines below the baseline. Human mobility calculated using the tweets number does not observe an upward trend after September. Google Community Mobility Report only found a marginally rebound less than a 5%, which is used to be 10% or more in other regions, so that the seasonal impact on human mobility in the Halton residential areas is too small to find easily.

Human mobility in the employment area rapidly decreases to below 50% of the baseline value during the first outbreak. This is followed by a quarterly change, rising between April and June, but then falling rapidly again in early July, with the process repeating again from July to September. This pattern is also reflected in the Google Community Mobility Report, but it is hard to detect. Currently, there is no information can explain this phenomenon yet. There is no significant trend is found for human mobility in open spaces calculated by the tweets number, but Google found a decrease during the first outbreak and seasonal impacts in the rest of the year. Human mobility derived from Twitter was broadly above the baseline value between May and October, and a significant increase was found in May, with a relatively significant downward trend also found after September. The reason for this phenomenon may be due to low data volume, which causes dramatically fluctuations. The human mobility change pattern in Halton utility and transportation areas is not similar to what is reported by Google, but rather closer to the trend in residential areas. Upon checking the map, the utility and transportation areas in land use maps are closer to residential areas. Therefore, the human mobility trends in utility and transportation areas are similar to the trend in residential areas.

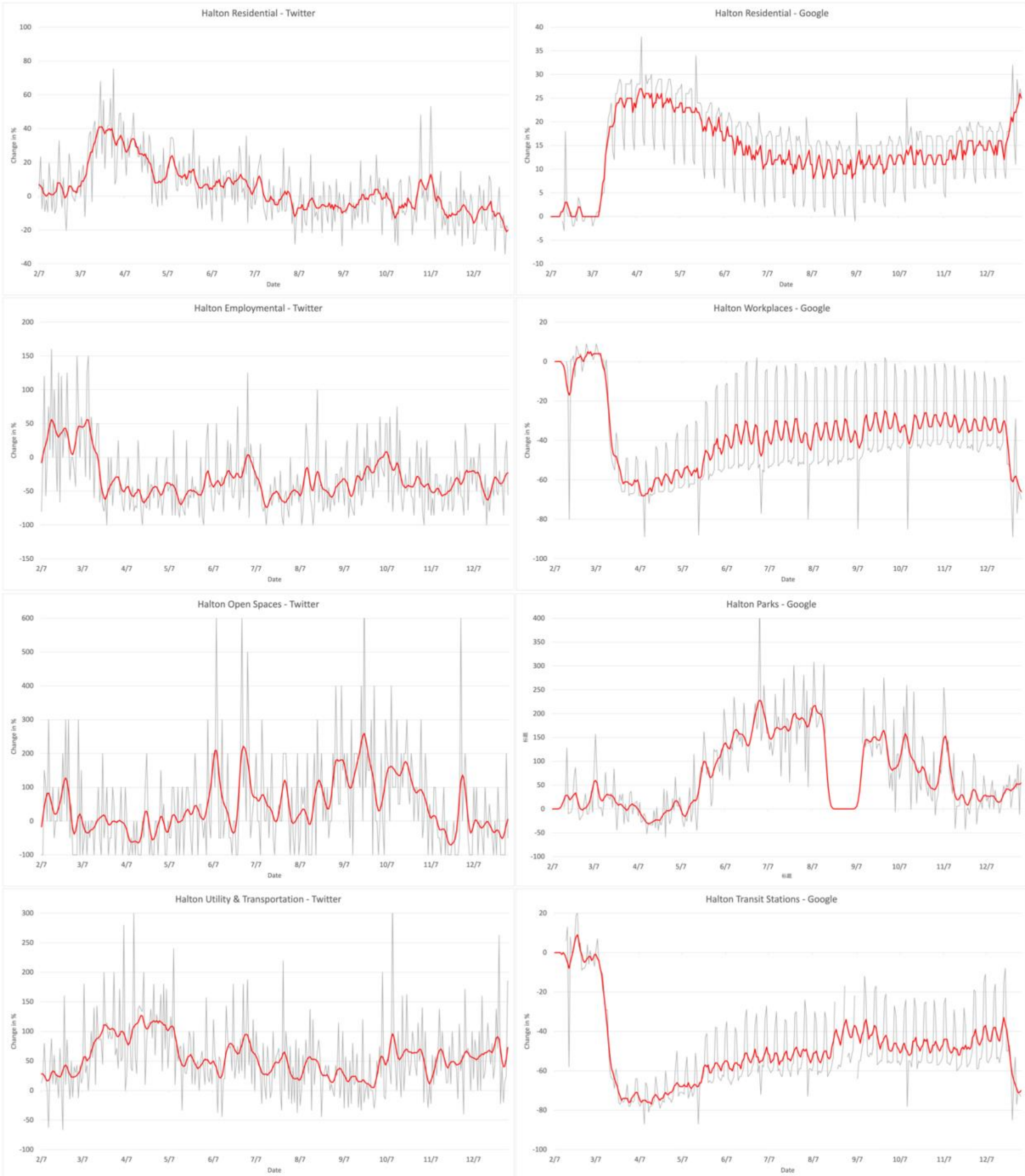


Figure 4.15 Daily human mobility change in four land use types in Halton Region.

4.3.5 Durham Region

Durham Region provides three land use types to verify, which are residential areas, employment areas, and utilities and transportation areas, respectively. Figure 4.16 presents the pattern of human mobility change calculated by the tweets number and the change pattern provided by Google, respectively.

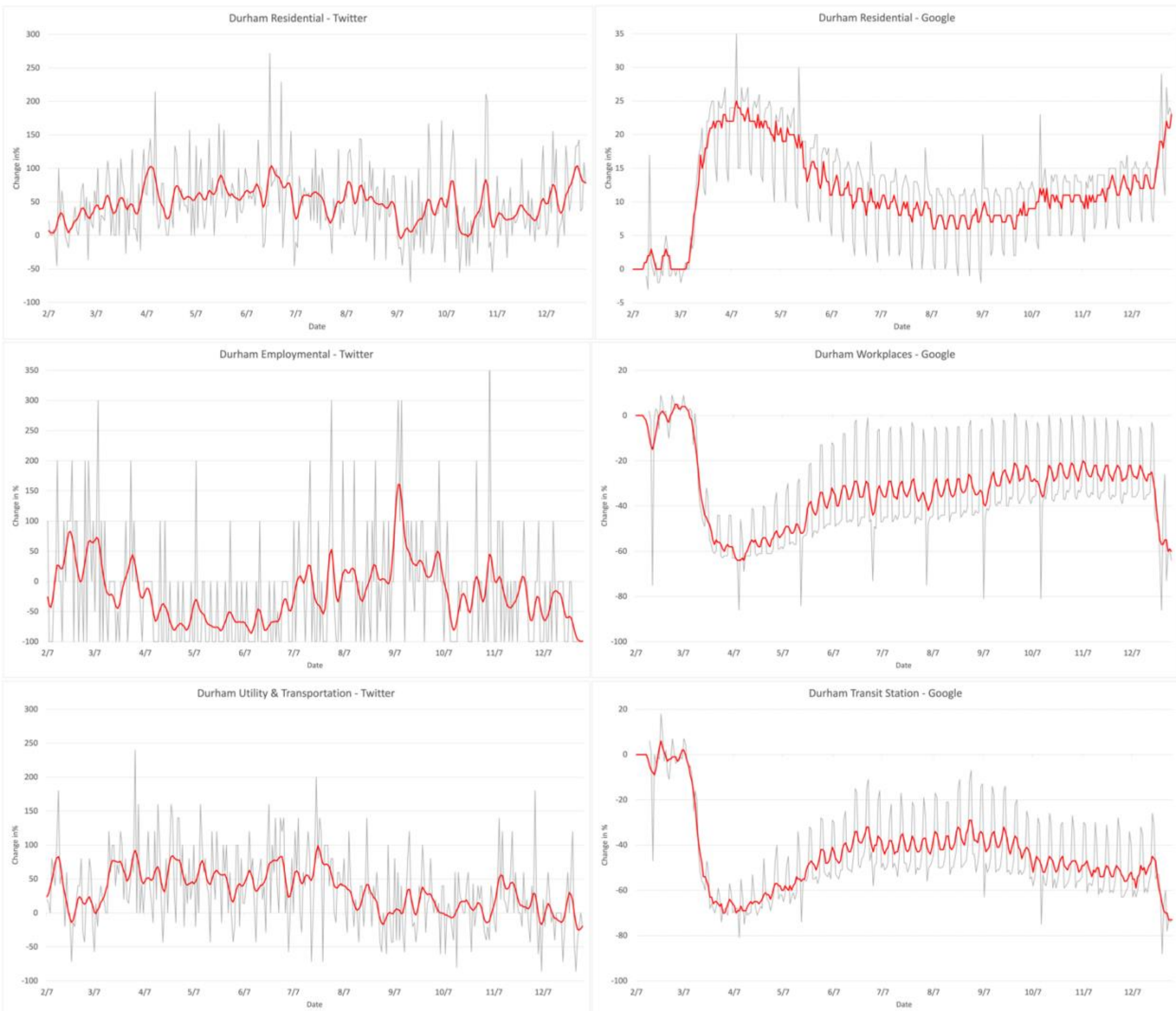


Figure 4.16 Daily human mobility change in three land use types in Durham Region.

The chart demonstrates that human mobility in the Durham residential areas remains at approximately 50% above the baseline value since the first outbreak until September. Also, a slight seasonal variation is observed with a slight decrease from June to August and a rebound from September to December. Human mobility in the employment area decreases to below 50% of the baseline value during the first outbreak while Google report shows only a 40% decrease in human mobility. Both Google-reported and twitter-derived human mobility patterns find an upward trend after June. However, Twitter-derived human mobility declines significantly after September while Google mostly remains constant. No significant trends are observed in human mobility in the utilities and transportation areas. Human mobility is primarily above the baseline value of approximately 50% in the first half of the year and falls back to above the baseline value in the second half of the year. Viewing the map of land use types identifies that Durham's utilities and transportation areas cover only a few of the city's road network and utilities. Therefore, the results produce a large inaccuracy with the Google Community Mobility Report.

4.4 Chapter Summary

This chapter addresses monthly and weekly spatial changes in human mobility based on land use polygons, and daily trends in human mobility based on land use types and dynamic baseline matrices. The roles are to explore how human mobility on land use polygons within the GTA is affected by pandemic severity, by policies, and to validate and identify issues with the Google Community Mobility Report, respectively.

Human mobility is found to be influenced by new confirmed cases in most areas in the monthly change patterns. Human mobility in residential areas increases significantly during the first outbreak. Then it decreases as the weather turned warmer and rebounds at the arrival of winter. Human mobility in some outdoor amenities declines during the first outbreak due to policy restrictions, but areas such as parks that remains open experiences a significant increase in human mobility. Human mobility in open spaces and rural areas show a significant seasonal increase from June to October. The qualities of open space and rural areas that allow for outdoor activities while keeping social distances make them be good destinations for trips during the pandemic. Human mobility in employment and institutional areas is typically lower than the baseline value during 2020 due to the shift to online teaching and work from home for all schools and non-essential businesses starting in late March. Thus, although the confirmed case number are relatively low during the summer, human mobility in employment areas and institutional areas are generally lower than before the pandemic. Human mobility in commercial areas is more sensitive to changes in the number of confirmed cases and policies, which receive a visible response almost within a week after release a new policy. Essential retail stores remain open even during the most severe period of the pandemic and provide curbside pickup and delivery services to stores. Therefore, the decline in human mobility in commercial areas is limited to a relatively small amount. The human mobility change in the utilities and transportation areas within the GTA depends mainly on what kind of land use types it connects to. Taking Highway 401 within the City of Toronto and the QEW within the Halton Region as examples, human mobility change on the road that connects open spaces and residential areas depend on the human mobility change in those areas. The number of commuters is significantly reduced in 2020, which leads to the human mobility in the utility and transportation areas is generally lower than the baseline value.

From the weekly study, most areas respond to the policy implementation within a week. It proves that most policies play a role in the fight against the pandemic, but some policies, such as limiting the number of people gathering, does not get a significant reflection in human mobility. While analyzing daily human mobility change patterns and comparing them with the Google Mobility Report, common patterns of mobility change are found for most land use types. Pearson correlation coefficients were attempted to be calculated to verify whether Twitter-derived human mobility obtained similar results to Google Community Mobility Report. Most land use types in 5 regions have a Pearson correlation coefficient between 0.1 to 0.4 where the p-value is less than 0.05, which means moderate positive correlations between Twitter- and Google-derived temporal human mobility patterns. Due to the sharing properties of tweets, the range of human mobility changes calculated using the tweets number is always smaller than that shown in the Google Community Mobility Report. In addition, differences in land use classification led to differences between the human mobility change patterns calculated by the tweets number and those provided by Google. Some additional information such as population and industrial structure are needed to explain those differences. Based on the charts of Google Mobility Report, differences between weekdays and weekends are more frequent. There are some spikes appear on public holidays, which means the human mobility during the public holidays is significantly different from the norm. However, those spikes impact the smoothed results. Comparison with the results of Jiang et al. (2021) found, the human mobility change patterns within the GTA using geotagged tweets are not identical to those found in New York City and are generally smaller in magnitude. These differences may be due to the demographic characteristics of New York are different from the GTA, and the division of land use types may also be different.

Chapter 5

Conclusions and Recommendations

5.1 Conclusions

In this thesis, the number of geotagged tweets is used to explore human mobility changing patterns based on either land use polygons or land use types in the GTA. Baseline data is selected from pre-pandemic period, which is January and February 2020. The mean and median are used to calculate monthly and weekly human mobility changes. A dynamic baseline matrix based on land use types is drawn upon to calculate the daily baseline values in order to facilitate comparison with the Google Community Mobility Report.

The analysis of monthly human mobility change patterns observed that the patterns are generally consistent with those found in previous studies. In the GTA, human mobility increases 20% rapidly then quickly fall down to pre-pandemic level during the first outbreak. Employment area and commercial area experiences a 30% and 10% decrease in human mobility, respectively. Human mobility in institutional areas and utility and transportation areas continuously decrease until 40% at the end of 2020. A special finding is that human mobility changes in open space and rural areas are also related to weather when there are no strict policies. In contrast to the results of human mobility in New York City, the human mobility patterns within the GTA and each region are similar, but rarely show a decline of more than 60%, even during the first outbreak, which is the most severely affected by the pandemic (Jiang et al., 2021). Despite some differences in the results, the epoch of this study provides a supplement to the human mobility change patterns during the second outbreak.

Most of the policies receive a relatively strong response in human mobility within one week. Human mobility in Weeks 1 to 10 responds quickly to new policy releases. The new policies issued for the second outbreak are relatively mild. As a result, it is difficult to observe significant changes in human mobility as reflected in these relatively mild policies. According to statistics provided by the Canadian government, the type of virus that caused the first two outbreaks are the same. The difference is that indoor contact rates may be the reason that cause the second outbreak being more severe than the first (Public Health Agency of Canada, 2022). Most policies for the second outbreak are also aimed at reducing indoor contacts, but they are too mild to control the transmission of the COVID-19 pandemic. Combining the human mobility changes in Stages 5 and 6, we can speculate that the more moderate policy did not strongly block the spread of the COVID-19 pandemic. Besides, the human mobility changes in the 7th stage observed a decrease in residential areas in some cities, which is conflict with theoretical results. Therefore, we assumed that the human mobility in Stage 7 not only influenced by policies but also by holidays.

The daily human mobility analysis provides human mobility change patterns across land use types within the GTA and each region. The first outbreak has a higher impact on human mobility for all land use types than the second. In general, human mobility changes in residential areas decreased after a rapid rise in the first outbreak. Human mobility in employment areas declines and then stays low. Human mobility in open spaces declines significantly during the first outbreak, but it is subsequently influenced by weather. Most human mobility trends calculated by the tweets number are similar to those provided by Google, but the magnitude is generally smaller than it provided by Google. The human mobility in utility and transportation areas is different from Google except the Peel Region, which may be due to the different land use classification. The closure of non-essential businesses or repetitive life routines may result in decreased sharing desire.

Overall, Twitter, a more accessible data source with high timeliness and few privacy concerns, obtains human mobility change patterns close to Google, especially in the early stages of the pandemic. In addition, human mobility derived using geo-tagged tweets can indicate the strictness and effectiveness of the policy. However, more variables should be comprehensively considered to determine whether the policy is applicable to the current situation. Some limitations to the present used approach will be described in detail in the next subsection, it makes a foreshadowing for crisis response and public health in future, especially for timely assessment of crisis and policy effectiveness.

5.2 Limitations

Although the research objectives of this thesis have all been achieved, there are still some limitations and space for improvement in both data and methods used at present, mainly in data quality and algorithm.

The first is that in cross-municipality studies, the quality of data provided by different municipalities varies greatly. Since the GTA is not an official authority, most of the land use data used in this study are provided by the government of each municipality. This makes the quality of land use data inconsistent. Some municipalities have not finish digitizing and interacting their land use data, the only way to collect data is to find land use data from higher-level administrative areas that cover the region, such as Caledon. This results in coarser land use data in a small part of cities, and detailed land use data in others. In addition, most cities are not merged their land use data to multi-polygon features, which allows researchers to observe more details such as human mobility increased in parks that remained open after the closure of provincial parks. A small number of cities, such as Vaughan and Richmond Hill, merged their land use data, so only the average human mobility change on that land use type can be observed.

Next limitation is because of the nature of social media data. Sampling bias in social media data are proposed in almost all studies that use social media data. Social media user groups usually bias toward to younger users. The detailed user ratio dependent on demographic variables such as age, income, ethnicity, and educational level. Some aging and marginalized groups use social media in a less frequency (Zhang et al., 2021). In Canada, young people aged from 18 to 24, and from 25 to 34 are two main Twitter user groups, 65% and 54% of these two age groups have Twitter accounts respectively. While older people have less penetration rate on Twitter, only 27% of people aged greater than 55 have a Twitter account. Annual household income between $60,000 and $99,999 are most likely to use Twitter, as well as people with a bachelor's degree or higher (49%) (Gruzd & Mai, 2020). It is difficult to recognize the real demands of aging and other marginal groups in crisis response. In addition to the sampling bias, the data volume from Twitter has also been questioned before. The reason is that the number of geotagged tweets is even less than 1% of the total tweets. However, due to the large base volume, the volume and velocity of management and storage become one of the challenges in processing big data, and collecting sub-sample data is a common method to optimize data collection strategies for social science research currently. Therefore, only collect geotagged tweets may be reasonable according to Sloan and Morgan (2015). Furthermore, the fundamentals of social media are about sharing, but people's desire to share will change over time. However, when working from home becomes the norm, people share in a lower frequency (Akan, 2022). Combined with the passage of the daily mobility analysis, the results of using the tweet count to calculate human mobility in the early stages of the pandemic are more similar to the Google Community Mobility Report.

For the existing algorithm for daily human mobility analysis, data for the dynamic baseline matrix are only taken from the five weeks between January and February 2020. It is not available to separately analyze the impact of human mobility by pandemic severity or seasonality for both

Google and Twitter yet. As the time period becomes longer, the classification of some locations may also change. The sampling range of the baseline matrix should also be updated in time to obtain more accurate results. These limitations may provide some ideas to subsequent works, it is hoped to improve the results in future.

5.3 Recommendations for Future Research

Based on the limitation mentioned in the previous subsection, future work can focus on improving the data quality and updating the algorithm. The data quality is not uniform across municipality as mentioned above. The heterogeneity of land use data across municipalities is mentioned above. To address this issue, future work can use a grid or hexagon network with pre-defined cell size to classify land use according to the main land use types within the cell. Different cell sizes can be applied to human mobility studies at different spatial scales. Since the same land use types are not merged, it is still possible to observe fine variations between different grids instead of assuming that human mobility is evenly distributed among a particular land use type.

For sampling bias in social media data between different groups, information can be collected through other channels as a supplement, such as using social media data commonly used by different ethnic groups, or encouraging users from different groups to collect accurate location data as volunteers, etc. Expanding data capacity from Twitter can extract textual information from tweet contents, for example, if a user posts a tweet about cooking, computer games, etc., we can determine whether the user is staying in a residential area based on the context. For the problem that share desire changes over time, the percentage of tweets with particular hashtags or contents in the total tweets can be calculated to track changes. More specifically, for example, when people just begin working from home, calculate the percentage of tweets with "work from home" or

"WFH" as part of the total tweets. When work from home becomes the norm, the percentage of tweets that mention work from home should calculated again. The difference between two ratios represents the change of work from home.

The existing algorithm using dynamic baseline matrix from January to February 2020 does not consider the seasonal change in human mobility. Therefore, the baseline data should be changed among the temporal scale to eliminate the seasonal impact on human mobility. For a longer study period, a year-on-year comparison may be useful to reduce seasonal effects. Even use a regression model calculated by the data from several past years to estimate the human mobility in 2020, then compare with the data during the pandemic to extract the effect due to the COVID-19 pandemic. If future works can obtain better results based on these points suggested above, a better understanding of the impact of pandemics on changes in human mobility and a more comprehensive assessment of the effectiveness of policies may be available, which may contribute to the application of VGI in epidemiology and other crisis response domains.

References

Abouzahra, M., & Tan, J. (2021). Twitter vs. Zika—The role of social media in epidemic outbreaks surveillance . *Health Policy and Technology, 10*(1), 174-181.

Abrams, D. S. (2021). COVID and crime: An early empirical look. *Journal of Public Economics, 194*, 104344-104344.

Akan, M. (2022, October 12). *Working from home in the media - A quick summary.* Retrieved April 2023, from https://www.dropbox.com/s/7qpsrbpgrn350pt/WFH_Oct_11_Mert.pdf?dl=0

Attard, M., Haklay, M., & Capineri, C. (2016). The potential of volunteered geographic information (VGI) in future transport systems. *Urban Planning, 1*(4).doi: https://doi.org/10.17645/up.v1i4.612

Bai, W., Sha, S., Cheung, T., Su, Z., Jackson, T., & Xiang, Y.-T. (2022). Optimizing the dynamic zero-COVID policy in China. *International Journal of Biological Science, 18*(14), 5314–5316.

Basiouka, S., Potsiou, C., & Bakogiannis, E. (2015). OpenStreetMap for cadastral purposes: An application using VGI for official processes in urban areas. *Survey Review - Directorate of Overseas Surveys, 47*(344), 333-341.

Bertuzzo, E., Mari, L., Pasetto, D., Miccoli, S., Casagrandi, R., Gatto, M., & Rinaldo, A. (2020). The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures. *Nature Communications, 11*(1), 4264-4264.

Bisanzio, D., Kraemer, M. U., Bogoch, I. I., Brewer, T., Brownstein, J. S., & Reithinger, R. (2020). Use of twitter social media activity as a proxy for human mobility to predict the spatiotemporal spread of COVID-19 at global scale. *Geospatial Health, 15*(1), doi: 10.4081/gh.2020.882.

Brail, S., & Kleinman, M. (2022). Impacts and implications for the post-COVID city: the case of Toronto. *Cambridge Journal of Regions, Economy and Society, 15*(3), 495–513.

Brown, G., McAlpine, C., Rhodes, J., Lunney, D., Goldingay, R., Fielding, K., . . . Vass, L. (2018). Assessing the validity of crowdsourced wildlife observations for conservation using public participatory mapping methods. *Biological Conservation, 227*, 141-151.

Cascetta, E. (1984). Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator. *Transportation Research. Part B: Methodological, 18*(4), 289-299.

Cascetta, E., & Nguyen, S. (1988). A unified framework for estimating or updating origin/destination matrices from traffic counts. *Transportation Research. Part B: Methodological, 22*(6), 437-455.

Cascetta, E., Inaudi, D., & Marquis, G. (1993). Dynamic estimators of origin-destination matrices using traffic counts. *Transportation Science, 27*(4), 363-373.

Chen, T., Hui, E., Wu, J., Lang, W., & Li, X. (2019). Identifying urban spatial structure and urban vibrancy in highly dense cities using georeferenced social media data. *Habitat International, 89*, 102005.

Chiesa, V., Antony, G., Wismar, M., & Rechel, B. (2021). COVID-19 pandemic: Health impact of staying at home, social distancing and 'lockdown' measures - A systematic review of systematic reviews. *Journal of Public Health, 43*(3), e462–e481.

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., . . . Longini Jr., I. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science, 368*(6489), 395-400.

Crawley, M. (2020, May 27). *COVID-19 in Ontario is now primarily a Toronto-area problem, figures show*. CBC News. Retrieved from https://www.cbc.ca/news/canada/toronto/covid19-ontario-gta-coronavirus-new-infections-1.5584693

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*, 1658-1658.

Díaz, C., Fossati, S., & Trajtenberg, N. (2022). Stay at home if you can: COVID-19 stay-at-home guidelines and local crime . *Journal of Empirical Legal Studies, 19*(4), 1067-1113.

Fang, H., Wang, L., & Yang, Y. (2020). Human mobility restrictions and the spread of the Novel Coronavirus (2019-nCoV) in China. *Journal of Public Economics, 191*, 104272.

Feick, R., & Roche, S. (2012). Understanding the value of VGI. *Crowdsourcing Geographic Knowledge* (pp. 15-29). Springer.

Ferster, C. J., Nelson, T., Robertson, C., & Feick, R. (2018). 1.04 - Current themes in volunteered geographic information. *Comprehensive Geographic Information Systems* (pp. 26-41). Elsevier.

Foodland Ontario. (n.d.). *Availability guide*. Retrieved Apr 1, 2023, from https://www.ontario.ca/foodland/page/availability-guide

Gesualdo, F., Stilo, G., D'Ambrosio, A., Carloni, E., Pandolfi, E., & Velardi, P. (2015). Can twitter be a source of information on allergy? Correlation of pollen counts with tweets reporting symptoms of allergic rhinoconjunctivitis and names of antihistamine drugs. *PloS ONE, 10*(7), e0133706-e0133706.

Gruzd & Mai. (2020). The state of social media in Canada 2020. Ryerson University Social Media Lab. Version 5. DOI: 10.5683/SP2/XIW8EW

Goodchild, M. F. (2007). Citizens as sensors: the world of volunteered geography. *GeoJournal,* *69*(4), 211-221.

Google. (2022). *COVID-19 community mobility reports*. Retrieved July 10, 2022, from https://www.google.com/covid19/mobility/

Goranson, C., Thihalolipavan, S., & Tada, N. d. (2013). VGI and public health: Possibilities and pitfalls. In *Crowdsourcing Geographic Knowledge* (pp. 329-340). Dordrecht: Springer.

Habibi, Z., Habibi, H., & Mohammadi, M. (2022). The potential impact of COVID-19 on the Chinese GDP, trade, and economy. *Economies, 10*(4), 73.

Hall, J. P., Kim, T. J., & Darter, M. I. (2000). Cost-benefit analysis of geographic information system implementation: Illinois Department of Transportation. *Transportation Research Record, 1719*, 219-226.

Haworth, B. T. (2018). Implications of volunteered geographic information for disaster management and GIScience: A more complex world of volunteered geography. *Annals of the American Association of Geographers, 108*(1), 226-240.

Haworth, B. T., Bruce, E., Whittaker, J., & Read, R. (2018). The good, the bad, and the uncertain: Contributions of volunteered geographic information to community disaster resilience. *Frontiers in Earth Science, 6*. doi: 10.3389/feart.2018.00183

Haworth, B., Whittaker, J., & Bruce, E. (2016). Assessing the application and value of participatory mapping for community bushfire preparation. *Applied Geography, 76*, 115-127.

Heo, S., Lim, C. C., & Bell, M. L. (2020). Relationships between local green space and human mobility patterns during COVID-19 for Maryland and California, USA. *Sustainability, 12*(22), 9401.

Hicks, A., Barclay, J., Chilvers, J., Armijos, M. T., Oven, K., Simmons, P., & Haklay, M. (2019). Global mapping of citizen science projects for disaster risk reduction. *Frontiers in Earth*

Science, 7, doi: 10.3389/feart.2019.00226.

Holguín-Veras, J., Jaller, M., Van Wassenhove, L. N., Pérez, N., & Wachtendorf, T. (2014). Material convergence: Important and understudied disaster phenomenon. *Natural Hazards Review, 15*(1), 1-12.

Hossain, L., Kam, D., Kong, F., Wigand, R. T., & Bossomaier, T. (2016). Social media in Ebola outbreak. *Epidemiology and Infection, 144*(10), 2136-2143.

Hu, S., Xiong, C., Yang, M., Younes, H., Luo, W., & Zhang, L. (2021). A big-data driven approach to analyzing and modeling human mobility trend under non-pharmaceutical interventions during COVID-19 pandemic. *Transportation Research. Part C, Emerging Technologies, 124*, 102955-102955.

Hu, T., Wang, S., She, B., Zhang, M., Huang, X., Cui, Y., . . . Li, Z. (2021). Human mobility data in the COVID-19 pandemic: characteristics, applications, and challenges. *International Journal of Digital Earth, 14*(9), 1126 - 1147.

Jacobs, C., & Zipf, A. (2017). Completeness of citizen science biodiversity data from a volunteered geographic information perspective. *Geo-spatial Information Science , 20*(1), 3-13.

Jiang, Y., Huang, X., & Li , Z. (2021). Spatiotemporal patterns of human mobility and its association with land use types during COVID-19 in New York City. *ISPRS International Journal of Geo-Information, 10*(5), 344.

Joshi, A., Sparks, R., Karimi, S., Yan, S.-L. J., Chughtai, A. A., Paris, C., . . . Forgoston, E. (2020). Automated monitoring of tweets for early detection of the 2014 Ebola epidemic. *PloS ONE, 15*(3), e0230322-e0230322.

Klar, B. (2022). Understanding changes to human mobility patterns in Ontario, Canada during the COVID-19 pandemic. *Electronic Thesis and Dissertation Repository, 8953*.

Lai, S., Ruktanonchai, N. W., Zhou, L., Prosper, O., Luo, W., Floyd, J. R., Wesolowski, A., Santillana, M., Zhang, C., Du, X., Yu, H., & Tatem, A. J. (2020). Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature*, 410-413.

Lei, L., Huang, X., Zhang, S., Yang, J., Yang, L., & Xu, M. (2020). Comparison of prevalence and associated factors of anxiety and depression among people affected by versus people unaffected by quarantine during the COVID-19 epidemic in southwestern China. *Medical Science Monitor*, e924609-e924609.

Li, Y., Li, M., Rice, M., Zhang, H., Sha, D., Li, M., . . . Yang, C. (2021). The impact of policy measures on human mobility, COVID-19 cases, and mortality in the US: A spatiotemporal perspective. *International Journal of Environmental Research and Public Health, 18*(3), 1-25.

Li, Z., Li, H., Zhang, X., & Zhao, C. (2021). Estimation of human mobility patterns for forecasting the early spread of disease. *Healthcare, 9*(9), 1224.

Liu, L., Wang, R., Guan, W. W., Bao, S., Yu, H., Fu, X., & Liu, H. (2022). Assessing reliability of Chinese geotagged social media data for spatiotemporal representation of human mobility. *ISPRS International Journal of Geo-information, 11*(2), 145.

Liu, T., Yang, Z., Zhao, Y., Wu, C., Zhou, Z., & Liu, Y. (2018). Temporal understanding of human mobility: A multi-time scale analysis. *PloS ONE, 13*(11), e0207697.

Longhorn, R. A., & Blakemore, M. J. (2008). *Geographic Information: Value, pricing, production, and consumption.* CRC Press.

Maher, M. (1983). Inferences on trip matrices from observations on link volumes: A Bayesian statistical approach. *Transportation Research. Part B: Methodological, 17*(6), 435-447.

Mancini, M. C. (2013). Geographical indications in Latin America value chains: A “branding from below" strategy or a mechanism excluding the poorest? *Journal of Rural Studies, 32*, 295-306.

McClellan, C., Ali, M. M., Mutter, R., Kroutil, L., & Landwehr, J. (2017). Using social media to monitor mental health discussions − Evidence from Twitter. *Journal of the American Medical Informatics Association, 24*(3), 496–502.

Mowat , D., & Raf, S. (2020). *COVID-19: impacts and opportunities*. Toronto: City of Toronto. Retrieved April 2023, from https://www.toronto.ca/legdocs/mmis/2020/ex/bgrd/backgroundfile-157346.pdf

Mozas-Calvache, A. T. (2016). Analysis of behaviour of vehicles using VGI data. *International Journal of Geographical Information Science, 30*(12), 2486-2505.

Nathan, M. (2021). The city and the virus . *Urban Studies,* 4209802110583.

Nguyen, T. D., Gupta, S., Andersen, M. S., Bento, A. I., Simon, K. I., & Wing, C. (2021). Impacts of state COVID-19 reopening policy on human mobility and mixing behavior. *Southern Economic Journal, 88*(2), 458-486.

Okan, C., Weidemann, G., & Bailey, P. E. (2021). "Alone, but not lonely": The impact of COVID-19 on older persons and the role of technology in staying connected. In Moustafa A. A (ed.) *Mental Health Effects of COVID-19*, pp. 117-131. London: Academic Press.

Orea, L., & Álvarez, I. C. (2022). How effective has the Spanish lockdown been to battle COVID-19? A spatial analysis of the coronavirus propagation across provinces. *Health Economics, 31*(1), 154-173.

Polyakova, M., Udalova, V., Kocks, G., Genadek, K., Finlay, K., & Finkelstein, A. N. (2021). Racial disparities in excess all-cause mortality during the early COVID-19 pandemic varied substantially across states. *Health Affairs, 40*(2), 307-316.

Public Health Agency of Canada. (2022). *Federal, provincial, territorial public health response plan for ongoing management of COVID-19.* Retrieved April 2023, from https://www.canada.ca/content/dam/phac-aspc/documents/services/diseases/2019-novel-coronavirus-infection/federal-provincial-territorial-public-health-response-plan-ongoing-management-covid-19/fpt-response-plan-eng.pdf

Public Health Ontario. (2020, December). *Daily epidemiologic summary COVID-19 in Ontario: January 15, 2020 to December 17, 2020*. Retrieved April 2023, from https://files.ontario.ca/moh-covid-19-report-en-2020-12-18.pdf

Public Health Ontario. (2020). *Epidemiologic summary COVID-19 in Ontario: January 15, 2020 to June 9, 2020.* Retrieved April 2023, from https://files.ontario.ca/moh-covid-19-report-en-2021-06-10.pdf

Public Health Ontario. (2020). *Epidemiological summary: COVID-19 in Ontario: January 15, 2020 to April 7, 2020* . Retrieved April 2023, from https://files.ontario.ca/moh-covid-19-report-en-2022-04-08.pdf

Rahman, M., & Thill, J.-C. (2022). Associations between COVID-19 pandemic, lockdown measures and human mobility: Longitudinal evidence from 86 countries. *International Journal of Environmental Research and Public Health, 19*(12), 7317.

Roche, S., Propeck-Zimmermann, E., & Mericskay, B. (2013). GeoWeb and crisis management: issues and perspectives of volunteered geographic information. *GeoJournal, 78*, 21-40.

Rodrigue, J.-P., Comtois, C., & Slack, B. (2006). *The geography of transport systems.* Abingdon, Oxon, England ; New York: Routledge.

Santamaria, C., Sermi, F., Spyratos, S., Iacus, S. M., Annunziato, A., Tarchi, D., & Vespe, M. (2020). Measuring the impact of COVID-19 confinement measures on human mobility using mobile positioning data. A European regional analysis. *Safety Science, 132*, 104925-104925.

Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The use of twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 Pandemic. *PLoS ONE, 6*(5), e19467-e19467.

Singh, R., Zhang, Y., & Wang, H. (2018). Exploring human mobility patterns in Melbourne using social media data. In *Australasian Database Conference on Databases Theory and Applications*, pp. 328-335.

Sloan, L., & Morgan, J. (2015). Who tweets with their location? Understanding the relationship between demographic characteristics and the use of geoservices and geotagging on twitter. *PLoS ONE, 10*(11), e0142209.

Song, J., Zhang, L., Qin, Z., & Ramli, M. A. (2022). Spatiotemporal evolving patterns of bike-share mobility networks and their associations with land-use conditions before and after the COVID-19 outbreak. *Physica A: Statistical Mechanics and its Applications, 592*, 126819.

Souza, R. C., Assunção, R. M., Oliveira, D. M., Neill, D. B., & Meira, W. (2019). Where did I get dengue? Detecting spatial clusters of infection risk with social network data. *Spatial and Spatio-temporal Epidemiology, 29*, 163-175.

Special Investigations Unit. (2020). *SIU director's report - case # 20-TCD-124.* Retrieved April 2023, from https://www.siu.on.ca/en/directors_report_details.php?drid=908

Statistics Canada. (2021, March 11). *COVID-19 in Canada: A one-year update on social and economic impacts.* Retrieved from https://www150.statcan.gc.ca/n1/pub/11-631-x/11-631-x2021001-eng.htm#a5

Statistics Canada. (2022, March 10). *COVID-19 in Canada: A two-year update on social and economic impacts.* Retrieved from https://www150.statcan.gc.ca/n1/pub/11-631-x/11-631-x2022001-eng.htm

Sy, K. T., Martinez, M. E., Rader, B., & White, L. F. (2021). Socioeconomic disparities in subway use and COVID-19 outcomes in New York City. *American Journal of Epidemiology, 190*(7), 1234-1242.

The Conference Board of Canada. (2020, March). *The impact of Toronto's financial sector*. Retrieved March 2023, from https://www.conferenceboard.ca/in-fact/the-impact-of-torontos-financial-sector/#:~:text=Toronto%20is%20ranked%20sixth%20globally%20in%20total%20bank%20assets%20by%20The%20Banker.

The New York Times. (2022, July). *How George Floyd died, and what happened next*. Retrieved April 2023, from https://www.nytimes.com/article/george-floyd.html

Tourism Economics. (n.d.). *Toronto's visitor economy: An economic catalyst for the city and the region.* Toronto. Retrieved April 2023, from https://s3.amazonaws.com/tourism-economics/craft/Latest-Research-Docs/Toronto-Visitor-Economy-final.pdf

Twenge, J. M., & Joiner, T. E. (2020). Mental distress among U.S. adults during the COVID-19 pandemic. *Journal of Clinical Psychology, 76*(12), 2170-2182.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146-1151.

Waring, M. E., Jake-Schoffman, D. E., Holovatska, M. M., Mejia, C., Williams, J. C., & Pagoto, S. L. (2018). Social media and obesity in adults: A review of recent research and future directions. *Current Diabetes Reports, 18*(6), 34-9.

Wei, P., Jin, C., & Xu, C. (2021). The influence of the COVID-19 pandemic on the imports and exports in China, Japan, and South Korea. *Frontiers in Public Health, 9*, 682693-682693.

Whittaker, J., McLennan, B., & Handmer, J. (2015). A review of informal volunteerism in emergencies and disasters: Definition, opportunities and challenges. *International Journal of Disaster Risk Reduction, 13*, 358-368.

Williamson, I., Rajabifard, A., & Holland, P. (2010). Spatially enabled society. *FIG Congress 2010, "Facing the Challenges – Building the Capacity".* Sydney.

World Health Organization. (2020). *WHO coronavirus (COVID-19) dashboard*. Retrieved March 1, 2023, from https://covid19.who.int/table

Wu, K., Ma, X., Liu, H., Zheng, J., Zhou, R., Yuan, Z., Huang, Z., Zhong, Q., Huang, Y., Zhang, Z., & Wu, X. (2022). Effects of different levels of non-pharmaceutical interventions on hand, foot and mouth disease in Guangzhou, China. *BMC Public Health*, 2398-2398.

Xiong, C., Hu, S., Yang, M., Younes, H., Luo, W., Ghader, S., & Zhang, L. (2020). Mobile device location data reveal human mobility response to state-level stay-at-home orders during the COVID-19 pandemic in the USA. *Journal of the Royal Society Interface, 17*(173), 20200344.

Xue, L., Jing, S., & Wang, H. (2021). Evaluating the impacts of non-pharmaceutical interventions on the transmission dynamics of COVID-19 in Canada based on mobile network. *PLoS ONE, 16*(12), e0261424.

Yabe, T., Tsubouchi, K., Fujiwara, N., Wada, T., Sekimoto, Y., & Ukkusuri, S. V. (2022). Non-compulsory measures sufficiently reduced human mobility in Tokyo during the COVID-19 epidemic. *Scientific Reports, 10*(1), 18053-18053.

Yang, C., Xiao, M., Ding, X., Tian, W., Zhai, Y., Chen, J., . . . Ye, X. (2019). Exploring human mobility patterns using geo-tagged social media data at the group level. *Journal of Spatial Science, 64*(2), 221-238.

Zhang, C., Yang, Y., & Mostafavi, A. (2021). Revealing unfairness in social media contributors' attention to vulnerable urban areas during disasters. *International Journal of Disaster Risk Reduction, 58*, 102160.

Zheng, Z., Xie, Z., Qin, Y., Wang, K., Yu, Y., & Fu, P. (2021). Exploring the influence of human mobility factors and spread prediction on early COVID-19 in the USA. *BMC Public Health, 21*(1), 615.