URBAN SPRAWL AND NEIGHBOURHOOD VITALITY IN TORONTO: A GIS AND REMOTE SENSING ANALYSIS

by

Sergio Barrios Jr.

A research paper

presented to Ryerson University

in partial fulfillment of the requirements for the degree of

Master of Spatial Analysis (M.S.A.)

Toronto, Ontario, Canada

AUTHOR'S DECLARATION

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

I authorize Ryerson University to lend this Research Paper to other institutions or individuals for purposes of scholarly research.

Sergio Barrios Jr.

ABSTRACT

Poverty is increasing within some neighbourhoods in the City of Toronto. In the past, suburban expansion was documented as having a strong inverse relationship with poverty in the inner city; as the city expanded outward, poverty increased. Recently, the priorities of the municipal government have shifted from developing former suburbs to a focus on improving public services in neighbourhoods now considered as "in distress".

Recent literature suggests that it may be possible to utilize remotely sensed imagery as a data proxy in social science investigations. The benefit is that satellite and aerial images can be acquired much more frequently than traditional data sources such as census information. This research determined the extent of urban growth from 1994 to 2005 in the heavily urbanized southern portion of the Toronto Census Metropolitan Area (CMA). In addition, it attempted to recognize a relationship between ortho-photos and census variables in a GIS environment. The hypothesis is that extracted image texture can be used as a surrogate indicator of neighbourhood vitality in the City of Toronto. The advantage for public officials and planners would be the ability to apply this relation to examine poverty/inequality issues more often. This would allow for action to be undertaken sooner. The results showed that it was not possible to relate image texture measures to variables from distressed neighbourhoods. This may be related to Toronto's underlying social complexities and changing urban structure.

ACKNOWLEDGEMENTS

I would first like to acknowledge Dr. Wayne Forsythe and Dr. Tony Hernandez for their combined assistance in formulating my major research topic. I appreciate the time they took out of their schedules to meet with me and respond to any questions I had to ensure that I was on the right path.

I would like to thank Dr. Forsythe as well as Dr. Andrew Millward, both whom supplied image data. In addition, I would like to thank Harvey Low and Mat Krepicz at the Social Policy Analysis and Research Unit, City of Toronto, for their data contributions.

I would also like to thank Noel Damba and Dan Jakubek at the Geospatial Map and Data Centre of the Ryerson University Library for their thoughts and timely collection and transfer of a large quantity of data for my analysis.

I would like to express my gratitude to my family for pushing me onward and making sure I did not distract myself for too long. And to my friends who made sure I was distracted when a break was warranted from the long hours of work that this paper provided.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

ARI – Image Channel Arithmetic

CHASS – Computing in the Humanities and Social Science

CMA – Census Metropolitan Area

DA – Dissemination Area

EA – Enumeration Area

ERDAS – Earth Resources Data Analysis System

ESRI – Environmental Systems Research Institute

GHK Consulting - Gilmore Hankey Kirke Consulting

GIS – Geographical Information Systems

GLCM – Grey Level Co-occurrence Matrix

KMO – Kaiser-Meyer-Olkin

Landsat TM – Landsat Thematic Mapper

LICO – Low Income Cut-off

MAUP - Modifiable Areal Unit Problem

NDVI – Normalized Difference Vegetation Index

PCA – Principal Component Analysis

PNs Excluded – Priority Neighbourhoods Excluded

SPAR – Social Policy and Research

SPSS - Statistical Package for the Social Sciences

Chapter 1: Introduction

An issue that has occupied sociologists is urban poverty/inequality brought about by segregation (Van Kempen, 1994). Studying urban poverty has been a tradition in North America long since the establishment of the Social Science Review which was founded in 1927. In its first years, one in five papers published by this journal wrote about poor relief systems or social insurance abroad, mostly in European countries. Over time poverty studies began focusing inward and less globally (Glennerster, 2000).

The amalgamated City of Toronto is the largest city in Canada and attracts many businesses with thriving markets in tourism, design and manufacturing, and entertainment (City of Toronto, 2008). Despite its prosperity, poverty in the City of Toronto has increased since 1981 when it was primarily concentrated in the old City of Toronto. The United Way released a report which states poverty has spread into the inner suburbs of the amalgamated City of Toronto, e.g. North York, Scarborough (United Way of Greater Toronto, 2004) (Fig.1). Some key findings were that the number of poor families has increased 69% (73,900 to 124,000) between 1981 and 2001 and the number of very high poverty neighbourhoods has increased by 43% since 1981 (3 to 7). Thirteen priority neighbourhoods in the City of Toronto now exist and are subject to investigation for better public service investments (The United Way, 2008).

Another report by the Centre for Urban and Community Studies (CUCS) at the University of Toronto studied the distribution of income and other data from the 1971 and 2001 censuses and grouped the city’s neighbourhoods based on whether average income in each neighbourhood increased, decreased, or was stable over the thirty year time period. Their results found that the city’s neighbourhoods are polarized by income

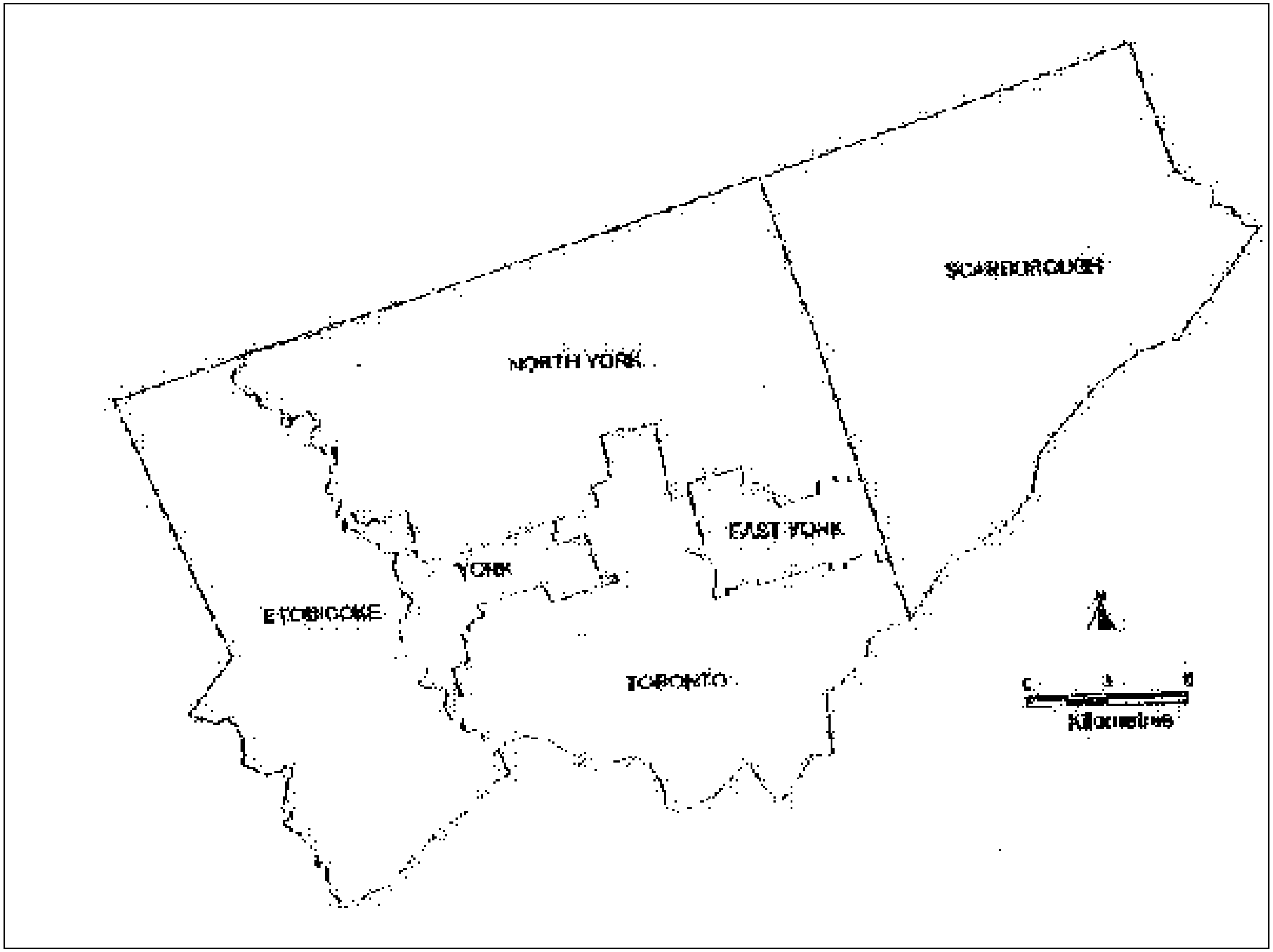

Figure 1.1 – Pre-amalgamated City of Toronto with former (inner) suburbs

and other ethno-cultural characteristics and the affluent and poor are concentrated in distinct areas (Centre for Urban and Community Studies, University of Toronto, 2007).

Policy makers and public officials, such as divisions within the municipal government of the City of Toronto, to a large degree depend on the Canadian Census as a source of social data to examine spatial trends, in a GIS environment, taking place within their city. Imperative actions needed by officials to allocate vital services in deteriorating neighbourhoods may be long over due with the release of census data every five years. An alternate source of data is needed that is not only released more frequently but is also applied in urban planning. Remotely sensed data has these qualities and is released to the public at no cost (depending on its source). It also has been used increasingly in urban environment analyses because they provide urban planners with crucial data needed for urban analyses (De Paul, 2007). Urban monitoring using remotely sensed data consists of 1) land use change detection (identifying type, amount and location of changes) and 2) land use impact analysis (evaluates the effects of such impacts to the environment). Studies such images, as part as Li and Weng (2007) and Jensen and Gatrell (2005) have reported on using remotely sensed data, extracted from of their social investigations in quality of life assessment and texture-census regression respectively.

Gentrification is driven by two mechanisms: by profit which may have unjust effects on the poor and working class, or to improve social infrastructure (Caulfield, 1994). Gentrification of poor neighbourhoods is a practice that has been applied in Toronto's neighbourhoods since the 1970's. This was brought about after the 1950s and 1960s when the municipal government made it a priority to expand the suburbs at the expense of inner city areas (Goliath, 2004). This type of urban development is called sprawl which can be defined as:

> a pattern of urban and metropolitan growth that reflects low-density, automobile dependent, exclusionary new development on the fringe of settled areas often surrounding a deteriorating city (Jargowsky, 2002).

A common trend found in major cities is the movement of the more affluent (or those who are able) to relocate to the suburbs to escape stresses found in the inner city, thus widening the gap between richer and poorer neighbourhoods (Heisz and Mcleod, 2004). Suburban sprawl could then be said to be a contributor to social segregation (Jargowsky, 2002). The shift of blame towards the suburban development for the woes found in the City of Toronto may have been warranted back when the emphasis by the municipal government was on suburban development. The emphasis, as mentioned, seems to have shifted more towards the re-development of areas within the City of Toronto. However, with the increasing poverty and social exclusion recently reported in the City Toronto, one may question how much sprawl in the inner suburbs is occurring and to what degree it affects inner city poverty today. Though the city may have a history of suburban sprawl related poverty (Squires, 2002), this trend may not be static through out time. With gentrification to improve social infrastructure occurring in poor neighbourhoods (for example the Regent Park Revitalization City Initiative (City of Toronto, 2007)) and other parts of the city, the attraction suburban sprawl may have in alluring residents away from the inner city may not be as influential in more recent times than it may have been almost 50 years ago. The post amalgamated City of Toronto may be facing different challenges that bring about poverty in its inner suburbs, such as low-income families moving to the inner suburbs to find affordable housing which is proving increasingly elusive (CTV.ca News Staff, 2004).

. Suburban sprawl can be detected with the aid of extracted image data, i.e. texture. Texture is associated with urban applications of change detection because it is able to distinguish urban features from other types well, thus raising the accuracy of land cover classifications (Forsythe and Waters, 2006; Hong et al., 2005). Since texture is already related to the urban landscape, if texture can also be related to urban poverty characteristics then perhaps varying values of texture may give insight to degrees of poverty within neighbourhoods. Since suburban development commonly starts out by building low-density, relatively more homogeneous (or low heterogeneous) areas, a hypothesis can be developed using texture. Areas of low heterogeneity (spaced out housing, large lawns/yards, less compaction of varying surfaces) can be associated with the once viewed, stress-free lifestyle of suburbs. Areas of higher heterogeneity (higher dense, small residential lots, compaction of varying surfaces) can be associated with a possibly more stressful lifestyle of the inner city. Public officials and planners can then benefit from analysing urban poverty trends using texture on more frequent basis. Thus, allowing them to plan and develop policies more timely. However, texture's relation to urban poverty has yet to be proven on a consistent basis, and for the City of Toronto.

1.1 Research Objectives

Though not the main objective, the first portion of this research will utilize texture (among a few other image transformations) to distinguish urban features and perform a change detection to isolate urban development that had occurred between 1994 and 2005 for the City of Toronto, within the Toronto CMA using Landsat images. This will give an idea as to how much new urban development has occurred within the inner suburbs of the amalgamated City of Toronto (in relation to urban development occurring in the rest of the CMA) and if this portion seems reasonable to blame for the increasing poverty/inequality currently in the city.

The main focus of this research is to reveal a relation between image texture and neighbourhood vitality, from a time series analysis, with the goal of designating texture as a surrogate indicator of poverty. This will be carried out by first creating mosaics of ortho-photos for the City of Toronto for 1997 and 2005. Had this research been able to create mosaics of the Toronto CMA, the spatial effects between the City of Toronto and its surroundings would have been examined since the two are inter-related, economically for example (e.g. job opportunities, affordable housing, people travelling to work).

Ortho-photos were used because of their very high resolution which will allow subtle urban features to be detected. Satellite images offer coarser resolutions that would offer lower precisions. Texture analysis will be performed on both mosaics to extract contrast texture values. The mean contrast values from the 1997 and 2005 mosaics will be calculated for census polygons for the City of Toronto for 1996 and 2006 respectively. After joining a set of neighbourhood vitality indicators to the census polygons for both years, regressions will be produced using the mean contrast values as the dependent variable. To give insight into the spatial differences within the City of Toronto, three study area extents will be examined through regression for each year, i.e. 1996 and 2006: 1) thirteen priority neighbourhoods, 2) all neighbourhoods in the City of Toronto excluding the thirteen priority neighbourhoods, and 3) and all neighbourhoods (City of

Toronto). Therefore, six regression models will be produced in total. The regression model results will demonstrate how significant the relationship is between areal texture and areal/neighbourhood vitality indicators for the City of Toronto.

Chapter 2: Literature Review

2.1. History between Urban Growth (Sprawl) and Poverty

Cities experience growth, commonly outwards and beyond the boundaries of the main core. This development is uneven because lower populations settle outside the inner city in the suburbs while higher populations are found in the inner city. Suburbs have evolved from the traditional dormitory communities to new industrial poles generating the "suburbanisation" of employment and living opportunities (Gilbert, 2004). The word "sprawl" has been used on occasion by scholars, public officials, and community organization leaders who tangle with the variety of tests brought by urban life (Squires, 2002). Among the issues caused by urban sprawl, aside from others such as concern for the environment, is that of poverty.

Sprawl, as Rivers (2004) states, is associated with and equally responsible for "broader environmental degradation and societal dislocation". Toronto first began with a core and suburbs (i.e., the suburbs which are now part of Toronto since amalgamation in 1998). After WWII, Toronto was considered the financial and industrial centre of Canada. It was described as a vibrant region with a safe and diverse urban core (Donald, 2002). However, in recent decades, poverty in countries across North America and Europe has become more concentrated and fixed in particular areas of large cities (Bradford, 2007), such as Toronto. Solomon (2007) gives a detailed historical account of sprawl for the City of Toronto beginning in the early 20th century. In short, Toronto started receiving large waves of immigrants around 1931. Crowded districts raised concerns of moral and public health. People were then encouraged to relocate to the outskirts of Toronto in fear of such conditions leading to deviant behaviour, a cause that would then lead to sloth and poverty, lack of hygiene and disease. Movements, such as *The Garden City Movement* in the U.K., seeded the notion of relocating and building in the suburbs of City of Toronto with visions of a clean, open environment and access to services but were aimed in particular to the more affluent (Solomon, 2007). But the movements were failures in that they did not pull enough people. Suburbs in the 1950's, like North York, had sub-standard living conditions and falling levels of education. The tax-base in these areas was not prepared to provide governments with funding for required services. Services and living environments may begin to decay reducing neighbourhood vitality as services migrate to the suburbs.

The appeal of living in suburbs in the past was separation from the possible challenges of city life such as poverty and racial conflict. Wealthier households in pursuit of such lifestyles moved to suburbs which in turn can caused middle-income families to then move from the centre of the city to take the place of the wealthier families, leaving behind low income families to represent a larger portion of the tax base (Squires, 2002). Neighbourhoods comprised of less-well-off households where low income rates exceed 40% are then considered low income neighbourhoods (Heisz and Mcleod, 2004).

Urban build-up, or sprawl, is relevant to the economic segregation and the concentration of poverty. As Jargowsky (2002) explains, the process is a spiral between the "pull" of suburban life that extracts people away from the city core and the "push" of people to the suburbs because of unfavourable neighbourhood attributes. The relative balance between the "pull" and "push" varies over time, so, as Jargowsky states, an argument is pointless as to which of the two forces drives the process of sprawl and urban decay.

Today urban sprawl or urban change has continued on the outskirts of the City of Toronto: building permits have risen hitting a high in 2005, and new housing construction was maintained within the recent decade (Statistics Canada, 2006). Yet poverty conditions are still present within the inner city and former suburbs to this day with high poverty neighbourhoods having increased as well (MacDonnell, 2007).

Despite the common trends in history of suburban sprawl related poverty in developed inner cities, movements have changed and priorities are now focusing on improving the inner city infrastructure for sustainable urban development (City of Toronto, 2008). Examples are community conversations where Torontonians may voice their opinion in public meetings about how poverty may be reduced. Migrations to the suburbs may not be as apparent anymore and quite possibly the opposite may be occurring. Crime in the City of Toronto has been reduced overall (with the exception of homicides) (Toronto Public Service, 2008), and cultural positives such as places of worship, learning academies and other meeting places have been sprouting in strip plazas for newly arrived Canadians (e.g. Wexford area in Scarborough). Therefore, it is possible that suburbs are not to be blamed for the inner city woes as Toronto's inter-complexities have changed in coping with changing economic realities, waves of new Torontonians, and an infrastructure from 50 years ago (Soknacki, 2007).

2.2 Neighbourhood Poverty in Toronto and Measuring Vitality

Poverty is an issue that has not been fully understood because of how it has been approached. Seen as a homogenous mass, populations living in poverty have been handled with standardized programs of assistance, intended to serve the public at coarser spatial levels than more localized ones. Mismatches between local needs and solutions are bound to occur (Kotler et al., 2006).

The amalgamated City of Toronto today consists of several smaller clusters of low income neighbourhoods that surround a much more affluent downtown; unlike other Canadian CMA's which have one low-income cluster in the downtown core (Heisz and Mcleod, 2004). The City of Toronto and The United Way have identified 13 priority areas (neighbourhoods) that require immediate attention. These areas have been targeted for social investment and support through various social programs (The United Way, 2008). They were identified by isolating neighbourhoods where there was a 20% larger than average population for particular social services, and neighbourhoods that were 20% worse than the amalgamated City of Toronto average on vitality indicators (Strong Neighbourhoods Task Force, 2005).

Statistics Canada has always resorted to low income measures as indicators to measure the extent of how not well off some Canadians are on the basis of income (Income Statistics Division, Statistics Canada, 2007). The low income cut-off (LICO), the most commonly accepted Canadian poverty measure, was developed over 40 years ago to measure relative income distribution (Chase, 1993). LICOs have been used by analysts who wanted to study the characteristics of the relatively worse off families in Canada (e.g. a social indicator, a characteristic of one's standard of living and relative place within the economy (Chase, 1993)). However, Statistics Canada clearly emphasized that they are quite different from measures of poverty and have not been endorsed by Statistics Canada as such (Statistics Canada, Household Surveys Division, 1998). Despite the low income cut-off being described as supporting the relative "notion" of poverty and a valid measure, it does not account for the cost of living. It also includes a variety of people, many who may not consider themselves as poor/worse off (Low, 2008).

The City of Toronto set out to explore a list of neighbourhood vitality indicators, through Gilmore Hankey Kirke (GHK) Consulting, and released a report in 2005 with a set of possible variables that could potentially be applied in neighbourhood condition investigations (Doblias and Battye, 2005). Formulated from similar indicator reports conducted in other parts of the world, the list composed for Toronto was created for its neighbourhoods by not only using the quantitative indicators used elsewhere, but also a number of qualitative measures that focus more precisely on the community. The indicators are lengthy to list and comprise several domains that encompass certain sections of vitality.

2.3 Social Science collaborating with Remote Sensing

Researchers have been merging social science data with remote sensing data in analyses for different purposes. Some researchers have explored how to effectively display areal census data spatially for visualization and analysis, e.g. Martin (1989) suggests interpolation from population weighted centroid locations (of census polygons) derived from a grid. This resulted in a raster where cells were assigned weights representing the probability of a cell receiving a portion of the current centroid's population. The dasymetric approach can be used to dis-aggregate census data at a fine spatial resolution by using remote sensing in a GIS environment (Chen, 2002) by essentially transforming socioeconomic data from arbitrary zones into a physical geography which can then allow correlations to be established between land divisions and zonal census data (Chen, 1998).

Other researchers have done much work to use census data as ancillary information for remote sensing image classification purposes. Different applications can be found in work by Hutchinson (1982), Sadler and Barnsley (1990), and Vogelmann et al. (1998) who utilized area-based census data where confusion between land classes could not be resolved initially by the classification algorithm, thus applying GIS data to assign classes to unresolved pixel clusters.

And other researchers focused on the scale effect and the zonal effect using area-based census data. Many data sets are available only at spatial levels of association that are lower than that in which they were collected. Aggregation (reducing a larger number of areal units to smaller number of areal units in a given area) changes the values of any statistical analysis for the variables in two different ways. The change in scale causes a loss of information due to fewer data values to work with. This is the scale effect. The second affect the size of the larger areal units which the smaller areal units will be aggregated to will affect any statistical analysis. This is the zonal effect. Both effects make up a geographic research concern called the modifiable areal unit problem (MAUP) which will cause statistical results to vary (Dark and Bram, 2007). Any attempts to apply results taken from lower levels (e.g. census tracts, enumeration areas) to higher levels (individuals, households) is known as ecological fallacy (Dudley, 1991). Flowerdrew and Green (1992) describe several methods that estimate values from a source data set to a target data set, with both sets differing in spatial resolution, e.g. areal interpolation based on weighting.

Though much work has been applied to joining zonal census data and pixel-based remote sensed data, little work has been done to identify correlations between the two sets of data (Chen, 2002). Wilkinson (1996) describes three main ways in which remote sensing and GIS data (including census data) have been used to enhance each other:

(1) remote sensing is used like a tool to collect data for use in GIS - detecting urban expansion (Cheng and Masser, 2003),

(2) GIS data used as ancillary data to improve remotely sensed derived products (as described earlier), and

(3) GIS and remote sensing are both used together for modeling and analysis – estimating population and residential density (Landford et al., 1991).

The integration of the census data and remote sensing data has also been used to identify poverty pockets, locate housing sites for low-incomers, and assess the quality of life (Lo and Faber, 1997).

Thus far, these methods have involved census data converted from vector to raster formats because of the similarity of remote sensing and raster GIS data models. Expansion in integrated GIS software packages allows the extraction of image data based on GIS polygons (Wicks et al., 2002).

2.3.1 Aggregating remotely sensed data to areal polygons

The conceptual work for integrating extracted image data with socioeconomic data was laid out by Chombart de Lauwe, a French sociologist who was interested in using aerial photography and developed the concept of "l'espace social" (social space) in 1952 (Lo and Faber, 1997). Chombart de Lauwe was known for developing this concept for quality of life assessment. Only two pieces of literature were found which integrated remotely sensed data to census polygons for analysis. Li and Weng (2007) studied the quality of life in Marion County (Indianapolis), Indiana, USA. They were able to aggregate remote sensing image characteristics to the block group level. The goal was to integrate a list of quality of life indicators comprised of derived census variables (e.g. population density, percentage of college above graduates, unemployment rate) and environmental variables (greenness, impervious surface and temperature from a Landsat ETM+ (Enhanced Thematic Mapper Plus) image) of the study area. Factor analysis was performed on the list of indicators to reduce the dimensionality and redundancy seen through high correlations between variables in a matrix as a preliminary analysis. They used factor analysis because it has the Kaiser-Meyer-Olkin, KMO, and Bartlett's test which display whether redundancy has been removed. Once redundancy is removed the data is suitable for factor analysis. They went on to create spatial indices each one characterizing an aspect of quality of life. The indices were based on the factor loadings of entered variables and amount of variance explained. The indices were then regressed back against their variables. Their results produced high values of coefficient of determination (R^2). The research demonstrates that remote sensing and social data (census data) were integrated for social science purposes. Integration does, however, imply converting the format of one data type so as to use it in an environment that is familiar to the other, e.g., the aggregation of remotely sensed data to census polygons, thus bring MAUP into the fold.

Jensen and Gatrell (2005) integrated remote sensing data with census data. Though their work was briefer than Li and Weng (2007) they too aggregated remote sensing data, in this case texture, to census block groups. They simply performed regression analysis and found that median income was positively correlated with texture for the city of Terre Haute, Indiana, U.S.A. They went on to say that differences in texture may provide a “starting point” for professionals to select areas in an urban environment for further analysis. They suggest that local government officials, planners, and others may one day be able to use such a geo-technical integration which fosters the development of new, novel, and effective classification systems. The integration also benefits from an advantage of remote sensing data, i.e. more frequently available image data. Though there may be other sources of information on residents, such as customer data and survey data, these may come at acquisition costs the researcher may not be able to afford (i.e., time to acquire customer data or draft, survey and analyze survey data). Indicators for Toronto may require information from sources not readily available (e.g. neighbourhood safety indicators such as crime statistics sourced from police services). Other information may come from standard 5 year released census data. Remote sensing platforms have high data acquisition frequencies that are referred to as temporal resolutions (e.g. Landsat Thematic Mapper (TM) 5 acquires an image over the same area every 16 days) (Jensen, 2000). Incorporating remote sensing in social science studies provides information for more frequent trend analysis.

2.4 Detecting Urban Change in Remote Sensing with Texture Properties

Change detection in remote sensing focuses on land cover variations occurring on the ground by analyzing multi-temporal remote-sensing images (Bruzzone and Bovolo, 2007). Fostered by sprawl, the City of Toronto has expanded in the CMA changing the face of natural environments. This creates a diverse landscape of anthropogenic and natural materials such as concrete, asphalt, metal, water, grass, trees, shrubbery, and soil (Jensen and Gatrell, 2005). Detecting this change can be approached using remote sensing by observing “temporal variation in spectral response involving situations where the spectral characteristics of cover types in a given location change over time" (Hoffer, 1978).

It has been argued that the best method to determine urban change is through image radiometry (Ridd and Liu, 1998). Methods of unsupervised classification or supervised classification have been utilized for such goals. Classification errors will however ripple throughout the analysis. Therefore, it is recommended that radiometry be directly used for efficiency, that is if the images acquired are year to year and from the same satellite platform (Forsythe, 2003).

New approaches have incorporated the use of image texture to improve results. Texture measures worked very well in distinguishing urban built-up areas, and between those urban and agricultural features that have similar spectral signatures when used in classification procedures (Forsythe and Waters, 2006). Texture also isolates areas differing in homogeneity/heterogeneity (PCI Inc., 2003). Texture has been used to detect urban development as seen in Forsythe (2003) who studied the Toronto CMA, as well as in Sun et al. (2007) who studied the Calgary CMA and found an increase in the overall classification accuracy.

Chapter 3: Study Area(s)

In order to exhibit sprawl for the City of Toronto the study area must consist of a large region, one that shows sprawl relative to a large urban base core such as a city. The Toronto CMA will serve as the first of two study areas in this paper to reveal sprawl that has occurred between 1994 and 2005. The second area will be the City of Toronto post amalgamation. It will be used to examine the relations between social indicators and remote sensing data (texture) in a GIS environment. The processing time and data storage for the second part of the analysis required that only the City of Toronto be examined.

3.1 Toronto CMA

The first study area consists of a portion of the Toronto CMA (Fig. 3.1). A CMA is made up of one or more adjacent municipalities centred on a large urban area (also known as an urban core, i.e. City of Toronto). The census population count must be at least 100,000 people to form a CMA (Statistics Canada, 2002). In Figure 3.1 the CMA is made up of almost a dozen municipalities (some in portions): Vaughan, Brampton, Mississauga, Milton, Oakville, Richmond Hill, Markham, Halton Hills, Pickering, King, Caledon, surrounding Toronto. Toronto has been the third fastest growing city in North America behind Los Angeles and Dallas (Lorinc, 2001).

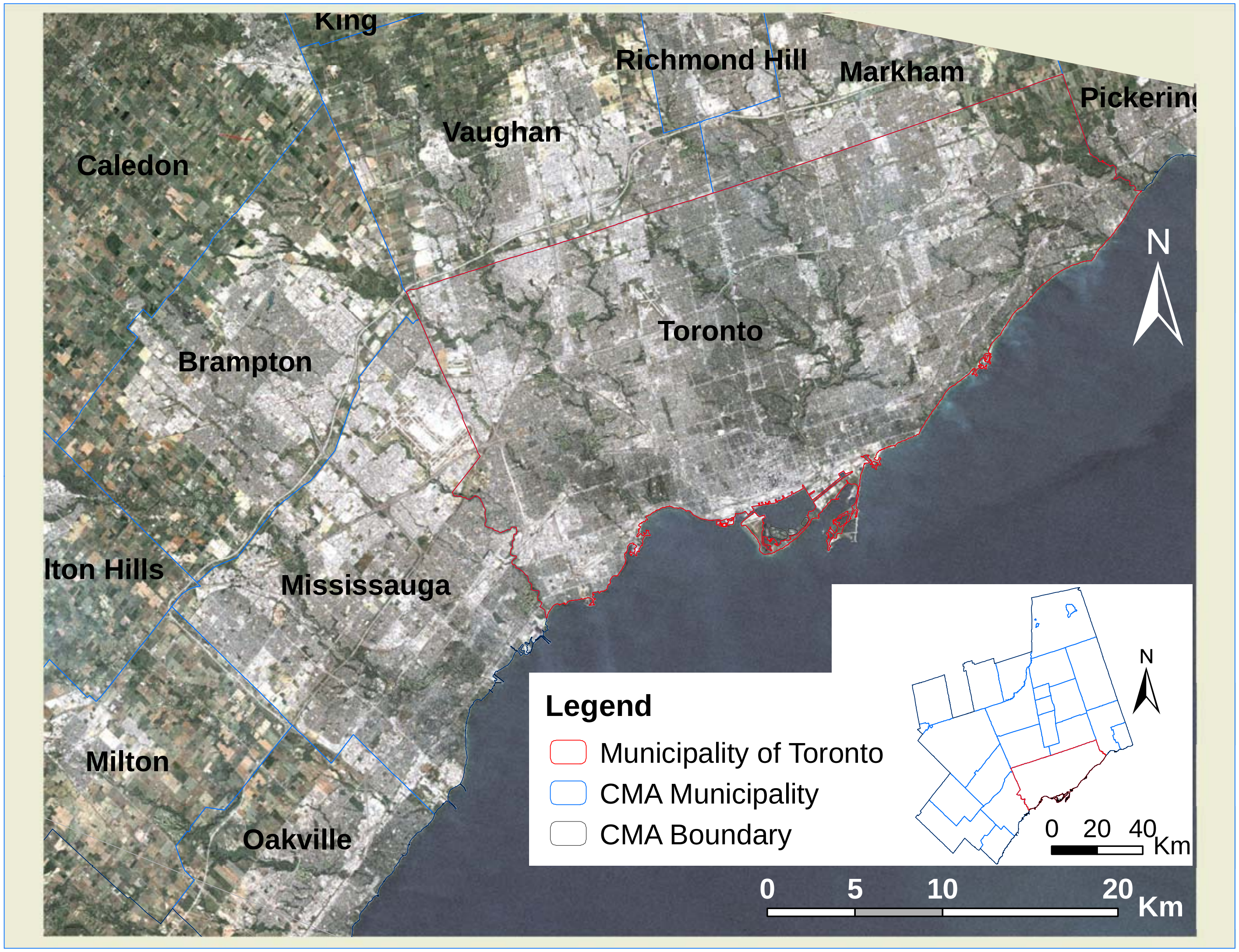

Figure 3.1 – Study Area Toronto CMA

3.2 City of Toronto

The second area is the City of Toronto consisting of neighbourhoods (Figure 3.2) that have been developed by the Social Policy Analysis and Research (SPAR) unit in the City's Social Development & Administration Division with assistance from Toronto Public Health. The boundary criteria for the neighbourhoods include:

1. a basis of an Urban Development Services Residential Communities map,
2. a make up of more than one census tract,
3. population of at least 7,000 – 10,000,
4. a joining of similar adjacent areas if first two criteria are met based on the population percentage of low income households,
5. respecting existing boundaries, and
6. preserving small neighbourhoods so services can include them in their service area, and "manageable" number of neighbourhoods for presentations/reports.

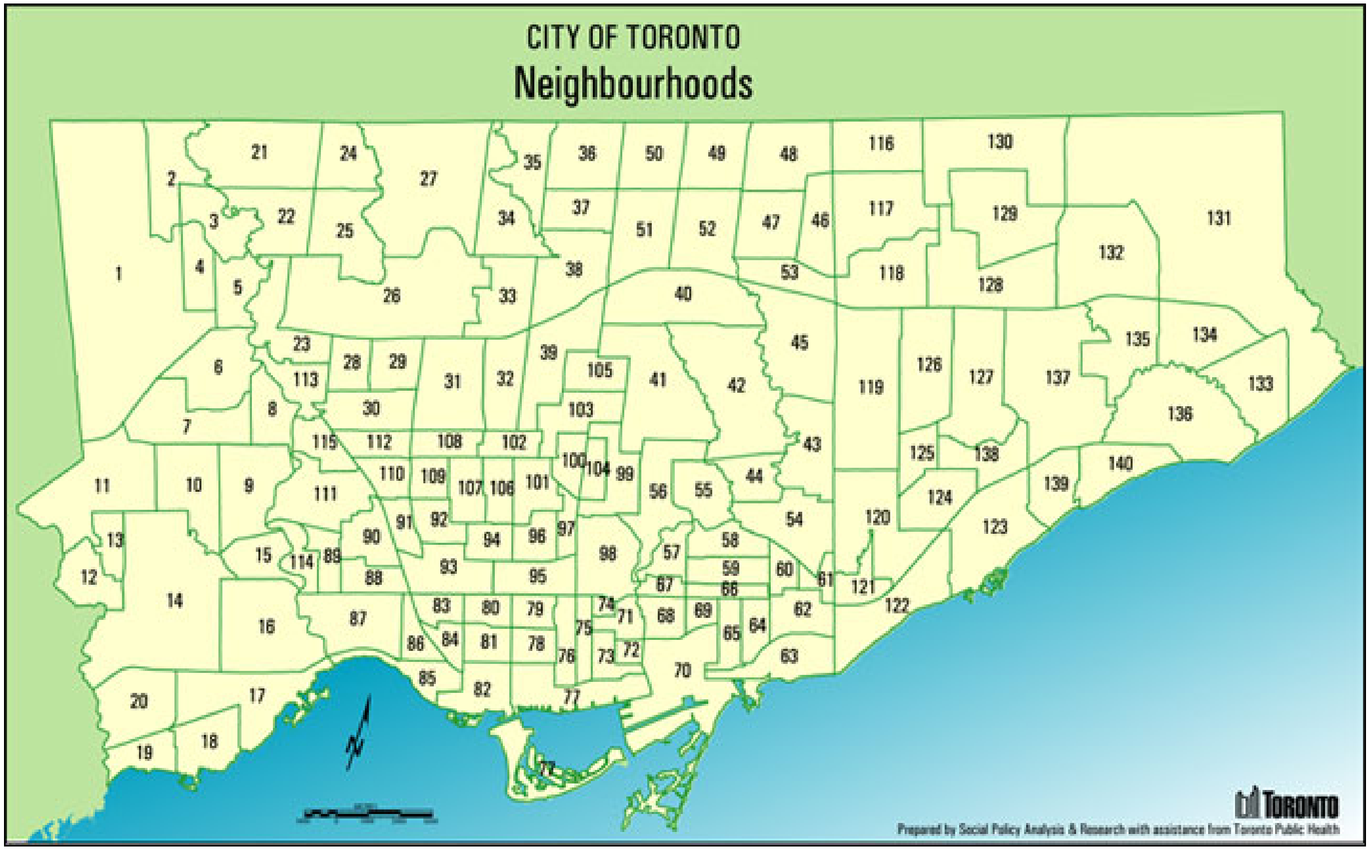

Figure 3.2 – Neighbourhoods of Toronto
Source: http://www.toronto.ca/demographics/profiles_map_and_index.htm

Chapter 4: Data Sets

The data collected for this report have been divided into part A and part B. Part A lists data used for urban development change detection between two Landsat TM images from 1994 and 2005. Part B will use aerial image data from 1997 and 2005 in combination with GIS data and census data to reveal any relations.

Part A

2 Landsat TM images: July 11, 1994 and July 25, 2005 (Toronto CMA)

- Resolution: 30 metre
- Spatial Reference: GRS 1980 Transverse Mercator
- Datum: GRS 1980
- Source(s): Professor Wayne Forsythe & Professor Andrew Millward (Ryerson University)

Part B

Ortho-photographs

(i) City of Toronto Ortho-photos 1997 (Land Information Toronto, 1997)
  - Resolution: 0.5 metre
  - Spatial Reference: UTM
  - Datum: 3 Degree

(ii) City of Toronto Ortho-photos 2005 (First Base Solutions, 2005)
  - Resolution: 0.2 metre
  - Spatial Reference: UTM Zone 17
  - Datum: Horizontal, NAD83

Shapefiles

(i) 1996 Enumeration Area Units of the City of Toronto (Statistics Canada, 1996)
  - Spatial Reference: GCS North American 1983
  - Datum: North American 1983

(iii) 2006 Dissemination Area Units of the City of Toronto (Statistics Canada, 2006)
  - Spatial Reference: GCS North American 1983
  - Datum: North American 1983

(iv) Priority Neighbourhood Areas (City of Toronto, 2005).
  - Spatial Reference: GCS North American 1927
  - Datum: North American 1927

Census Data: 1996 and 2006 at Enumeration Area (EA) and Dissemination Area (DA) level respectively (Canadian Socio-Economic Information Management System (CANSIM), 2007)

Chapter 5: Methodology & Results (Part A)

5.1 Data Processing: Noise/Haze Reduction Image Fusion, & Transformations

Before performing the unsupervised classification, the 1994 and 2005 images were first sub-setted so that they both shared similar extents and covered the same geographic area of the Toronto CMA.

5.1.1 Noise/Haze Reduction

The 2005 image, since it was to be classified, was processed to reduce both noise and haze interference. Haze is made up of trace gases and particle matter. Toronto occasionally experiences a thin layer of brown haze in the summer months, usually in July and August with conditions of low wind speeds and temperature inversion (Megaw et al., 1974). Aside from haze being hazardous to health, it also scatters radiation which can severely reduce the information content in an image until a point where the image loses contrast and it becomes hard to distinguish one object from another (Jensen, 2000). Earth Resources Data Analysis System (ERDAS) Imagine software carries both haze and noise reduction tools. The Haze Reduction tool reduces the overall haze using a Tasseled Cap transformation, yielding and removing a component correlated with haze. The Noise Reduction tool reduces noise while preserving the subtle details in an image.

Transformations of the 2005 image were processed to extract spectral information from the image that would display certain land features better for classification. The extracted spectral information can be interpreted as "features", remote sensing scene objects with similar characteristics. "Feature Extraction" encompasses a broad range of techniques, including some widely used techniques as the Normalized Difference Vegetation Index (NDVI), Principal Component Analysis, and Texture Analysis. Derived information from such techniques can be used as inputs to more advanced feature extraction techniques such as feature classification (Anderson, 2008).

5.1.2 NDVI

The first transformation performed was a Normalized Difference Vegetation Index (NDVI). First developed by Rouse et al. (1974), NDVI enhances the spectral differences between vegetation that strongly reflects near-infrared and the chlorophyll-absorption (red) part of the spectrum (Singh, 1989). NDVI has been proven to be a good detector of vegetation changes (PCI Inc., 2003). Where healthy, abundant vegetation is present, NDVI index values will be high (white) versus low values (black) where unhealthy or sparse vegetation and surfaces with no vegetation (such as bare soil, water, snow, ice or clouds) occur which can have near zero or slightly negative values (Mkhabela, et al., 2005). Forsythe and Waters (2006) found that NDVI worked well in segregating urban industrial/manufacturing and newly excavated areas from residential districts.

5.1.3 NDVI Results

The resulting NDVI product can be seen in Figure 5.1 which shows the presence of healthy, dense vegetation as white areas and as white “vein” like features. Black features on the outskirts of Toronto are cleared vegetation sites.

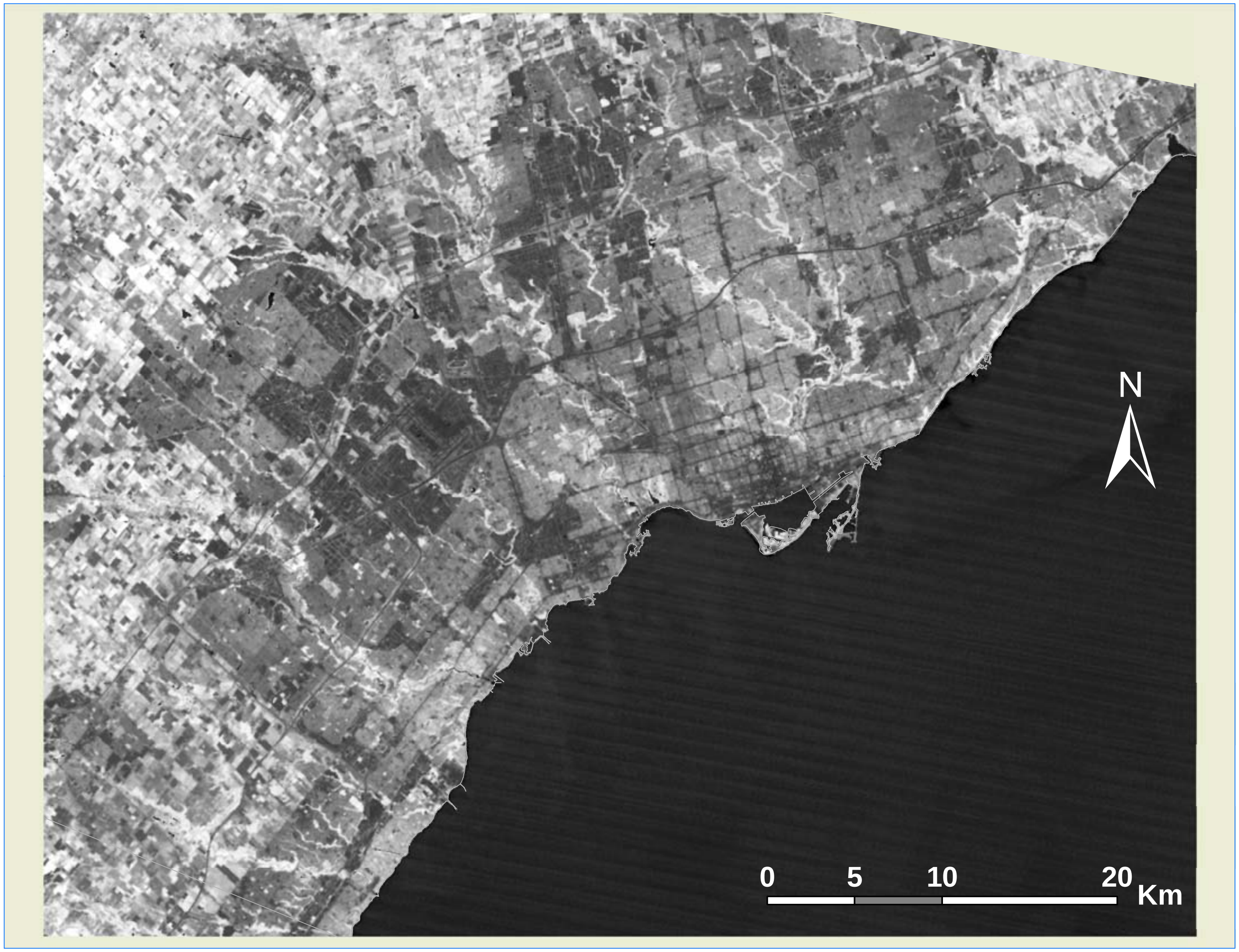

Figure 5.1 – Normalized difference vegetation index

5.1.4 PCA

The next transformation used was Principal Component Analysis (PCA). PCA reduces the dimensions of information within the original image to components (or eigenchannels) that explain image variance. PCA essentially takes image data and "packs" them together with the first few layers explaining most of the variation (PCI Inc., 2003). Lui and Lathrop, Jr. (2002) chose PCA for its simplicity, but PCA has been shown to be efficient in experimental results, especially when the number of principal components retained is properly selected (Azimi-Sadjadi et al., 1993; Benediktsson and Sveinsson, 1997). Forysthe and Waters (2006) found that PCA was effective in distinguishing urban green-space like parks and golf courses.

5.1.5 PCA Results

The components that were chosen are seen in Figure 5.2 and 5.3. Components 1 and 2 contain 99.18% (Table 5.1) of the variance when summed.

Table 5.1 Variances of the principal components (eigenchannels)

Eigenchannel	Eigenvalue	% Variance			
1	3274.1323	90.21	+		
2	325.5867	8.97	+	=	**99.18%**
3	22.3186	0.61	+		
4	5.7477	0.16	+		
5	0.9473	0.03	+	=	**0.81%**
6	0.5258	0.01	+		
Total	6642.61				

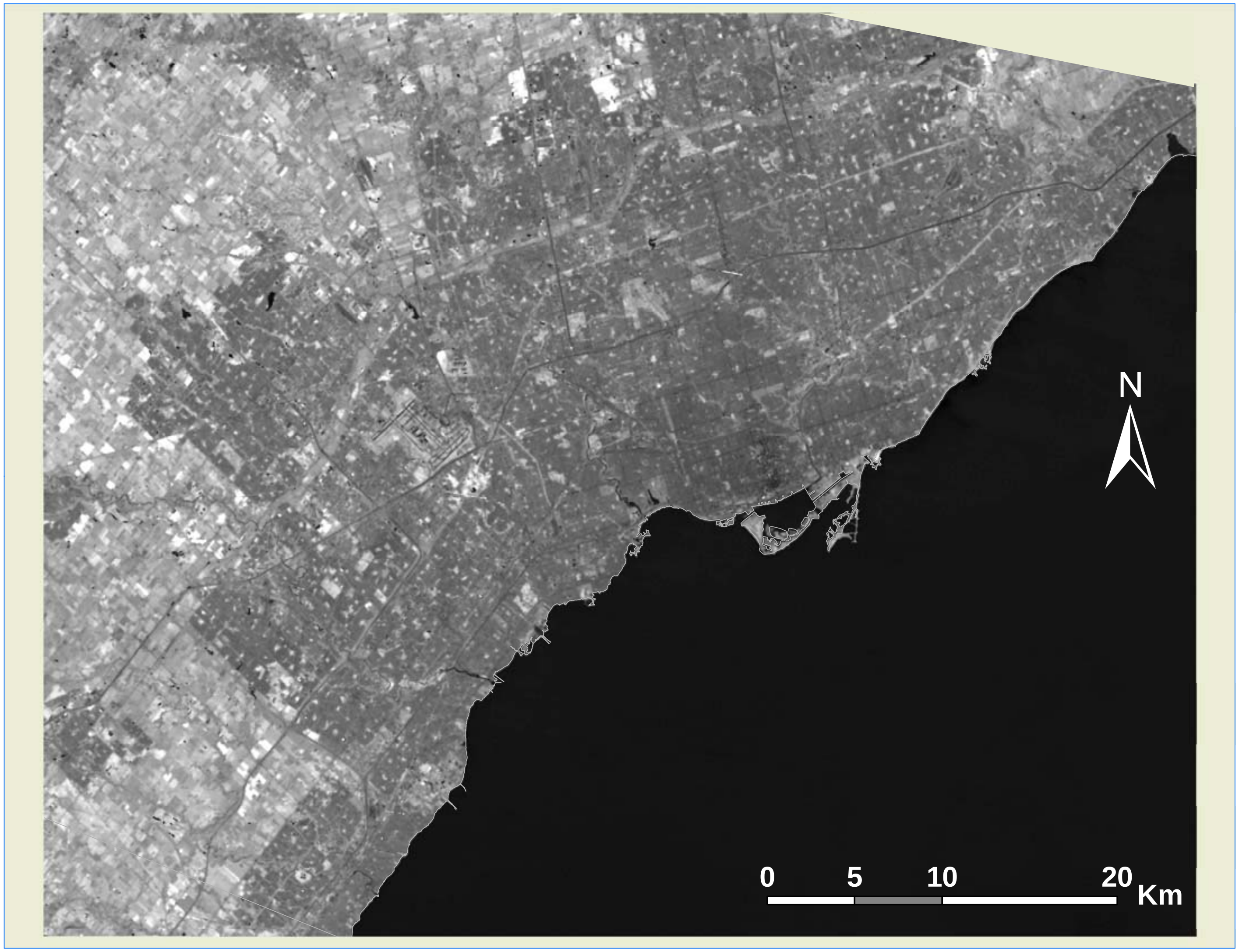

Figure 5.2 – Principal component 1

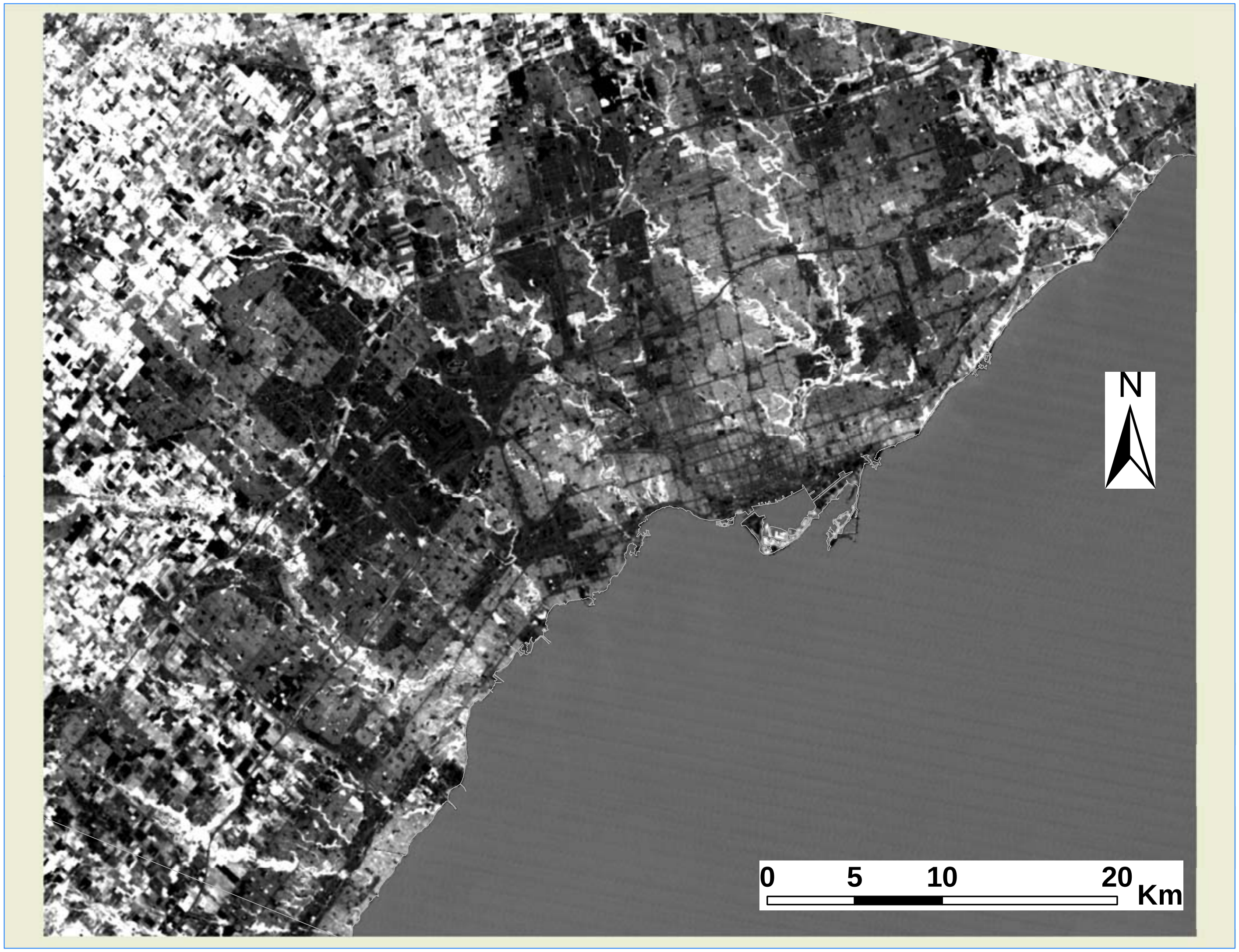

Figure 5.3 – Principal component 2

Component 1 (Fig. 5.2) distinguishes where major land cover types meet. For example, the medium grey area that encompasses the City of Toronto and adjacent areas can represent areas of heavily developed urban areas. The light grey areas to the North may represent areas that are not developed and vegetation/agricultural areas. Water is easily distinguishable by black. Component 2 (Fig. 5.3) depicts greater detail by displaying vegetation in white which makes fairly distinguishable not only in the rural areas in the North, but within the heavily development urban areas of the municipalities. Different urban development is also distinguishable from black to light grey.

5.1.6 Texture Analysis

The third and final information generation process was image texture analysis. It is known that the inclusion of texture measures raises classification accuracy (Forsythe and Waters, 2006). Texture is a characteristic that is an essential part of image data and usually important for target discrimination (Woodcock and Strahler, 1987). Texture analyzes grey level patterns and variations in a pixel's neighbourhood (Armenakis et al., 2003). Texture in an image implies that features in the image are not spatially homogenous, and produces an impression of roughness or smoothness created by tonal variation or the repetition of visual patterns across a surface (Irons and Petersen, 1981). Thus, non-homogeneity, or heterogeneity, of the features produces texture characteristics of different classes. Texture analysis would then be a logical step for image classification for urban/suburban and forest environments (Shaban and Dikshit, 2001).

Among the various texture algorithms, the grey level co-occurrence (GLCM) matrix approach is the most popular and effective (Conners and Harlow, 1980). GLCM is one of several statistical approaches to calculate texture in an image. To perform GLCM, a moving window is usually used to define a pixel’s neighbourhood with the texture quantified assigned to a centre pixel. GLCM’s window results in a matrix which approximates the joint probability distribution of a pair of grey levels that are separated at a certain distance and certain orientation (PCI Inc., 2003).

The window size can vary depending on the image’s resolution and the land feature in question. It is said that the size of the window, however, must be large enough to cover a texture pattern, but not too small so as to capture more than one (Pesaresi, 2000; Pusissant et al., 2005 as found in Forsythe and Waters, 2006). Since some researchers have now found that higher spatial resolutions are better suited with smaller window sizes (Zhang et al., 2003), a 3x3 and 7x7 were both tested. The 3x3 window showed more detailed results than the 7x7 and therefore the 3x3 size was used.

Band 2 was utilized in this study for texture analysis because, as mentioned in Forsythe (2004), band 2 produced similar results to previous studies which made band choice recommendations. Band 2 also seemed to distinguish developed urban areas from green-space areas.

5.1.7 Texture Analysis Results

Homogeneity texture (Fig. 5.4) and Mean texture (Fig. 5.5) were chosen among a number of texture measures offered in the PCI software package based on their ability to differentiate various land features from each other. Homogeneity in Figure 5.4 is fairly discrete in value with red circled areas as examples of high homogeneity (PCI Inc., 2003). A hypothesis was mentioned in the Introduction of the possible relation between high homogeneous-stress-free lifestyle of suburbs, and low homogeneous-stress apparent lifestyle of the inner city. This would suggest, if tested and proven, that the values depicted by the red circles indicate areas of more stress-free environments for residents compared to areas of lower homogeneity. However, these results would have had to been derived from higher resolution images (such as ortho-photos) to capture and distinguish subtle urban features in order to measure texture at a higher level to test this hypothesis accurately.

Mean texture measures the average grey level in a local window (PCI Inc., 2003). Mean texture in Figure 5.5 displays urban built-up areas in detail, as well as highlighting urban features and agricultural areas that have similar spectral signatures. Zhang et al. (2003) found in their experimental results using 10m panchromatic SPOT imagery that Homogeneity and Mean were among the top five pair combinations for classification accuracy.

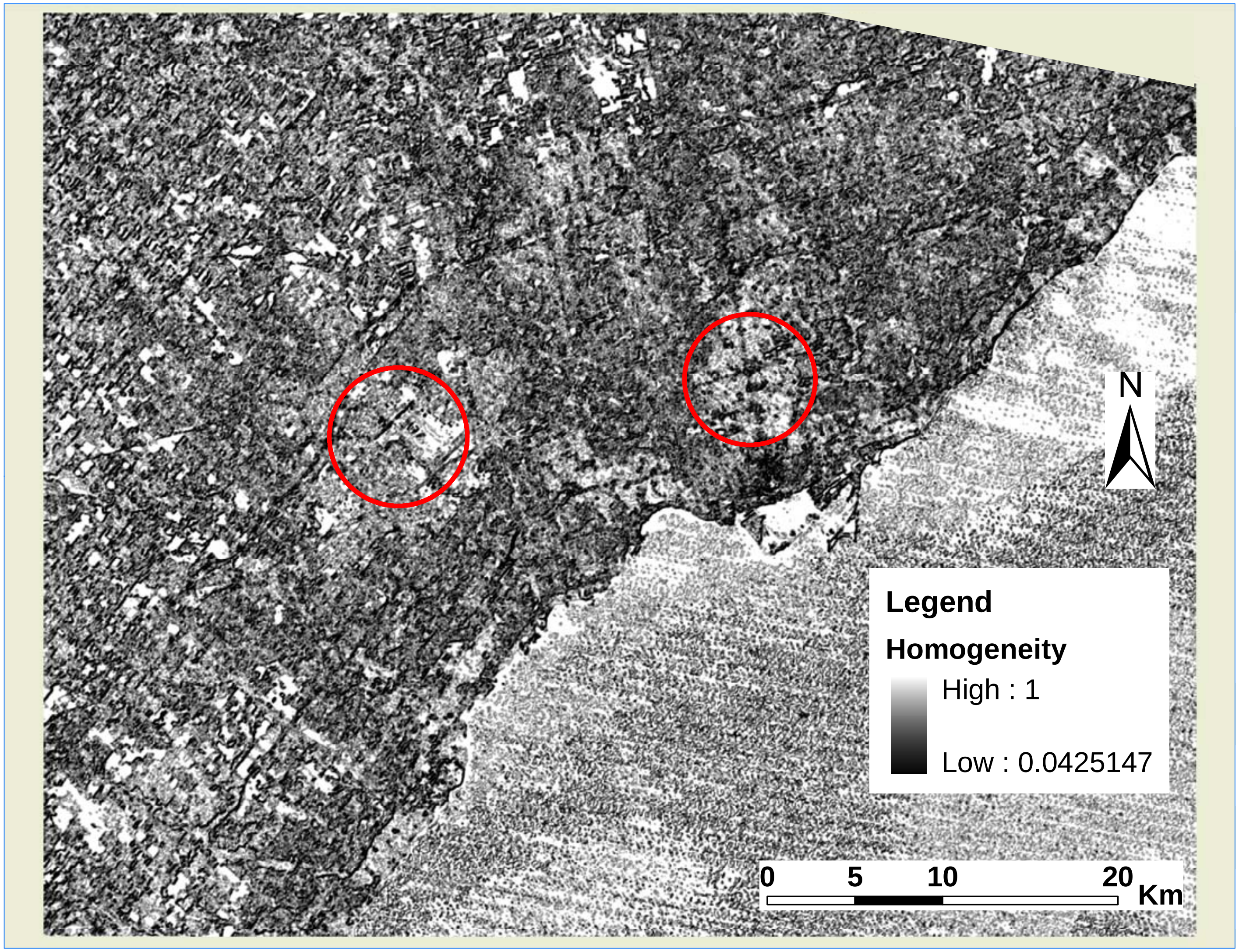

Figure 5.4 – Homogeneity Texture (red circles indicate example areas of high homogeneity)

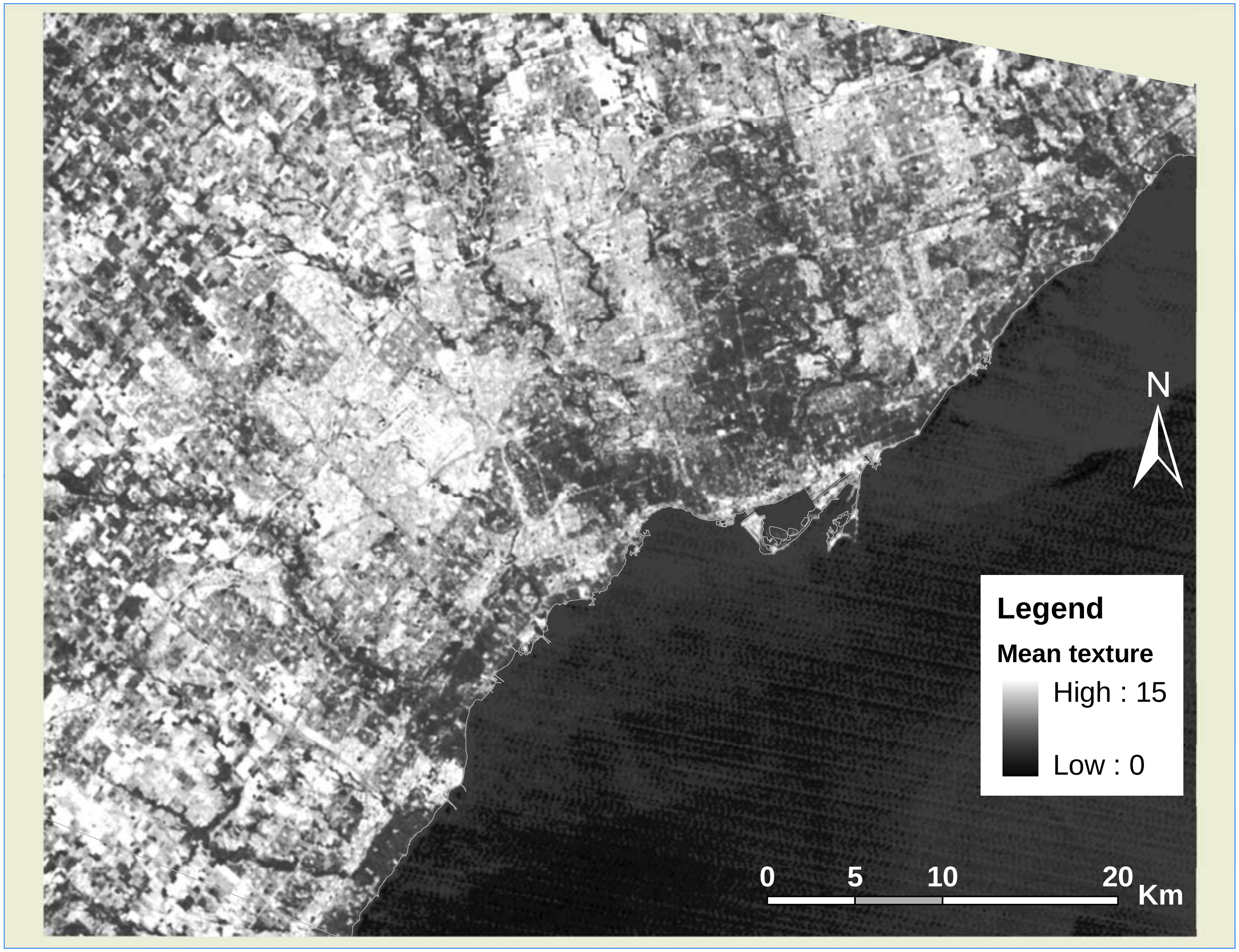

Figure 5.5 – Mean Texture

5.1.8 Stacked Layers

Next, nine layers were stacked to proceed with the classification process. The chosen layers began with bands 1 to 3 of the 2005 image. These three bands were used later on as the reference image in the accuracy assessment. Other bands, such as band 4 which distinguishes vegetation well, were not included because it may have been repetitive since band 4 was used in the NDVI calculation. Bands 5 to 7 are best known for uses not particularly vital in distinguishing urban features for this area of the Toronto CMA. The NDVI layer was added, as well as the first 2 components from PCA, and the homogeneity and mean texture measures. An empty band was also added as the input layer for the classification results in PCI.

5.2 Classification

The 2005 Landsat image was classified with unsupervised K-means clustering. The maximum number of clusters allowed in PCI was used in order to distinguish between classes in areas of recently ploughed fields that could be misinterpreted as land changed by recent urban development (Forsythe, 2004). Having used supervised classification using reference data (e.g. GPS points) this misinterpretation may have occurred due to ploughed fields being made up of similar materials as some urban areas, thus having similar spectral signatures. Using the maximum number of 255 would increase the odds of the clustering algorithm separating these two types of classes from each other before class aggregation. The 255 clusters were then aggregated into 4 general classes: water, urban build-up, green-space (i.e., forest, agriculture, tall grass, brush, and parks), and bare land. The water class may have been unnecessary and had it not been classified, the accuracy assessment would have only focused on how well land covers were classified.

Because Lake Ontario occupies 35% of the image much of the stratifying in the accuracy assessment would focus on how well water was classified. And since Lake Ontario is the most influential, if not the only, body of water visible from the image it may be pointless to include it in the classification since it is urban areas under examination. However, excluding Lake Ontario from classification requires that it be cut from the image. Had this been realized prior to classification a different accuracy assessment may have been acquired.

5.2.1 Classification Results

Figure 5.6 shows the results of the unsupervised classification with the 4 aggregated classes. .

5.2.2 Accuracy Assessment

The last step in the classification process was the accuracy assessment. This was performed in PCI where 197 random points were generated across the reference image which is displayed over the layer consisting of the 255 assigned classes. The reference image used here was the original 2005 image in true colour. Other options include using ancillary data (e.g. a map, other GIS data). But attaining ancillary data for the same year that the classified image was taken proved difficult to attain. Each point was then attributed to one of the four aggregated classes by zooming in and judging which class the pixel (that the point was located on) was assigned to. An error matrix plainly describes the accuracy of each aggregated class with errors of inclusion and errors of exclusion. The error matrix is then used to calculate a series of descriptive analytic statistics (Congalton, 1991).

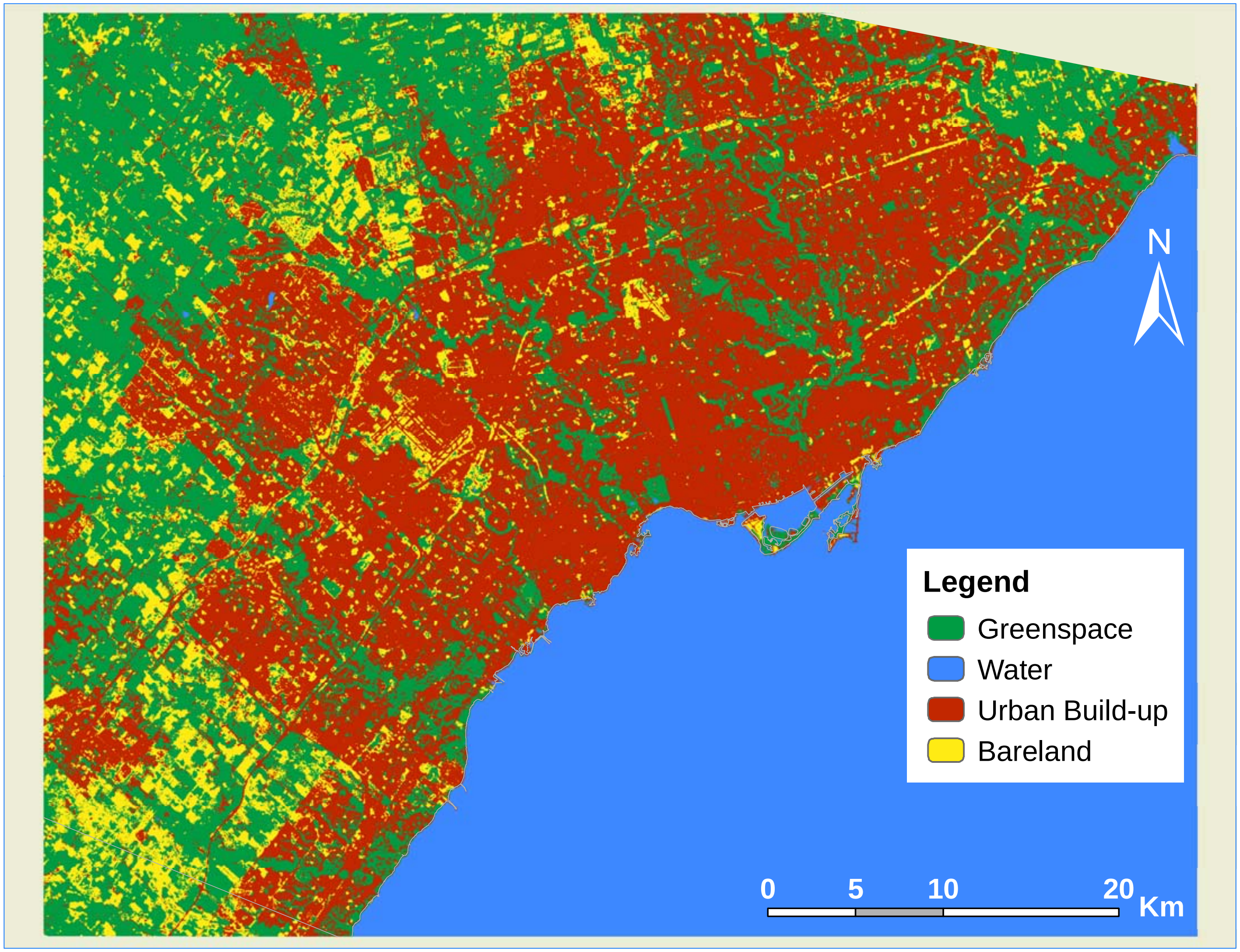

Figure 5.6 – Classification results with four aggregated classes

5.2.3 Accuracy Assessment Results

Table 5.2 and 5.3 give the results of the accuracy assessment in accuracy statistics and an error matrix respectively.

Table 5.2 Accuracy Statistics

Class Name	**Producer's Accuracy**	**95% Confidence Interval**	**User's Accuracy**	**95% Confidence Interval**	**Kappa Statistic**
Water	100.00%	(99.324% 100.676%)	98.67%	(95.404% 101.929%)	0.9788
Urban Build-up	82.86%	(73.314% 92.400%)	86.57%	(77.655% 95.479%)	0.7933
Greenspace	70.83%	(56.933% 84.734%)	77.273%	(63.754% 90.792%)	0.7010
Bareland	60.000%	(7.059% 112.941%)	27.273%	(-3.592% 58.137%)	0.2541

Overall Accuracy : 86.0% 95% Confidence Interval : (80.941% 91.059%)
Overall Kappa Statistic : 0.797% Overall Kappa Variance : 0.001%

Table 5.3 Error Matrix

Classified Data	Reference Data				
	Water	Urban Build-up	Green-space	Bare land	**Totals**
Water	74	1	0	0	**75**
Urban Build-up	0	58	8	1	**67**
Green-space	0	9	34	1	**44**
Bare land	0	2	6	3	**11**
Totals	**74**	**70**	**48**	**5**	**197**

Producer’s accuracy is the number of correct points that were assigned correctly to a class divided by the total number of points found that were randomly located in the class on the reference data (Story and Congalton, 1986), i.e. it indicates how accurately the analyst classified the image data by class column (NOAA, 2008). The user’s accuracy is the division of the number of random points correctly assigned to a class by the total amount of pixels that were actually classified to the same class, i.e., it measures how well the classification performed in the field by class row (NOAA, 2008). For example, for Urban Build-up although 82.86% was correctly identified as such, 86.57% of the area classified as Urban Build-up is actually urban build-up. User's accuracy allows the accuracy assessment to be stratified based on the areas of the classes. As mentioned before, water found in Lake Ontario occupied approximately more than 1/3 of the area found in the image. Thus, it is not surprising that 38% (74) of the randomly located points in the accuracy assessment were found in the Water class.

The producer's accuracy was high for water (100%) and urban build-up (82.67%). Water was fairly easy to classify as much of it was found in Lake Ontario (user's accuracy of 98.67%). Streams and rivers were a little difficult to classify, as well as differentiating water from shadow since both have similar spectral characteristics (explaining the 1 point that was assigned as Water but was classified as Urban Build-up).

Urban build-up involved classifying all types of urban features (high rise, industrial, commercial, and residential areas). Outside the downtown core of the City of Toronto are areas of urban build-up that are surrounded by bare land and vegetation. The hardship in classifying urban features in a city like Toronto was to not include features other than urban build-up. Toronto has many mixed landscapes close to each other including vegetated areas. Thus, the producer's accuracy (82.86%) for Urban Build-up seems understandable as it is not very accurate in the classification of a diverse environment. The user's accuracy indicates that 86.57% of the pixels classed as urban are urban on the ground.

Bare land (including ploughed land) was to be separated from other land disturbed by urban development by using 255 classes before class aggregation; unsupervised classification was necessary in order to create a mask to minimize this potential classification error (Forsythe, 2004). However, bare land was not easily distinguishable as there was confusion with urban and vegetated areas resulting in a low producer's accuracy (60.0%) and user's accuracy (27.273%).

Green-space included natural vegetation and agricultural vegetation. However, parts of these areas are close to urban areas which may cause confusion explaining why the user's accuracy (77.273%) is not high. The error matrix reveals the confusion in the classification process. Overall, the accuracy of the classification is 86.0% which is relatively good compared to a similar study of Toronto between 1999 and 2002 in Forsythe (2004) in which almost 96% was reached for overall accuracy using pan-sharpened Landsat ETM+ imagery for the sake of increased detail in the scenery. The overall accuracy is computed by dividing the total number of points from the major diagonal (i.e. total correct) by the total number of random points in the error matrix (Congalton, 1991). In a lecture on March 3, 2008, Dr. Wayne Forsythe stated that a minimum overall accuracy of 85% is needed today (W. Forsythe, personal communication, March 3, 2008), thus deeming this classification as acceptable.

5.3 Image band differencing and Change detection

5.3.1 Image band differencing

Image band differencing is the process of one dated image being subtracted from another dated image to produce a new image (Jensen, 1996). The 1994 image was subtracted from the 2005 image using the image channel arithmetic algorithm (ARI) in PCI to extract land feature change (white) that occurred between 1994 and 2005 (Fig.

5.7). Combining image band differencing and unsupervised classification to detect urban change eliminates errors that could occur if agricultural change is mistaken for urban change (Forsythe, 2003).

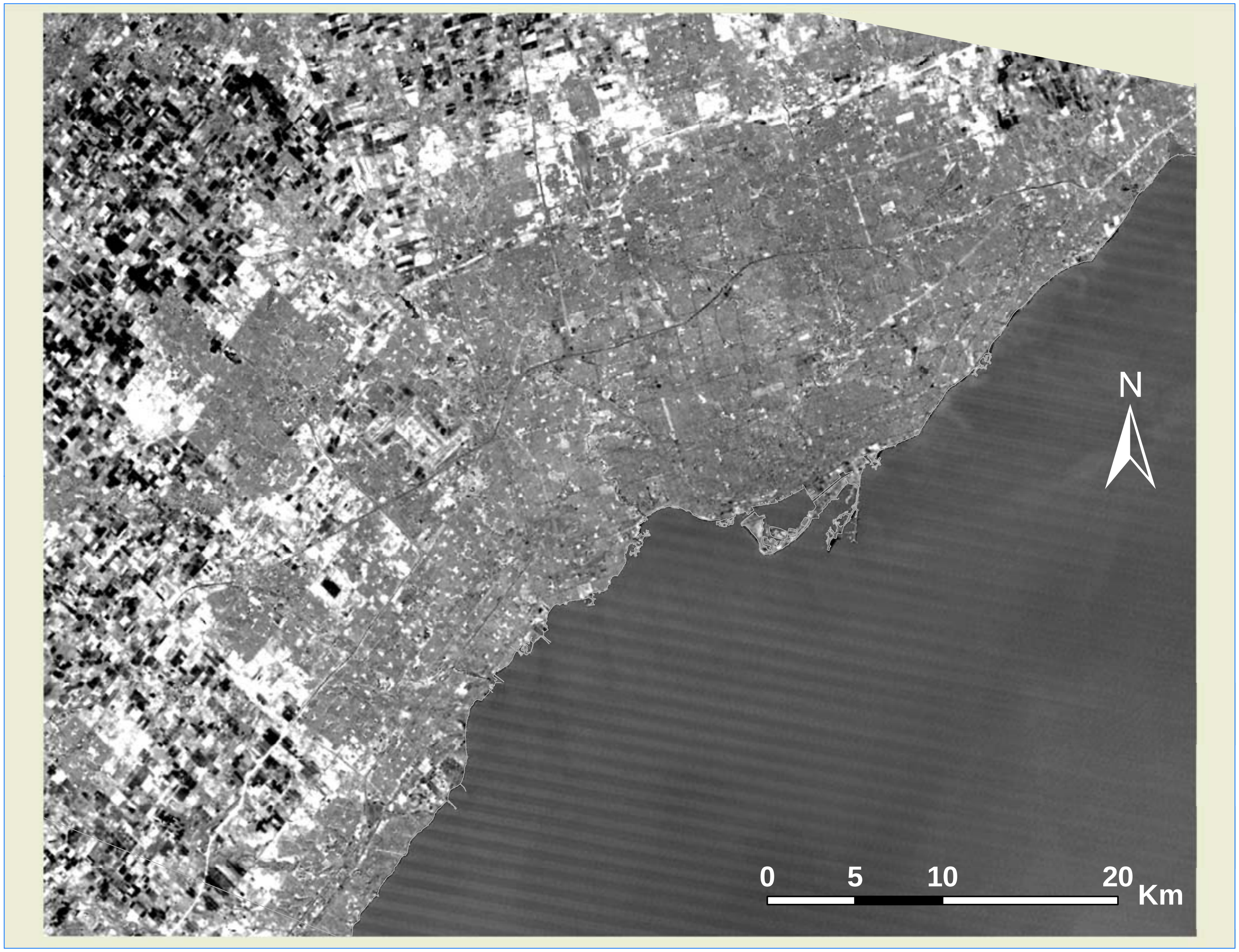

Figure 5.7 – Difference image showing extracted land feature change in white.

5.3.2 Class Aggregation and Image Calculation

In Environmental Systems Research Institute's (ESRI) ArcMap the classified 2005 image and the difference image were reclassified. The difference image was reclassified into two classes, one representing overall change between 1994 and 2005 and the other no change. The aggregated classified image was also reclassified 1 to 4. The reclassified difference image was added to the reclassified aggregated classed image using the Raster Calculator. The calculation gave eight classes, four which represented areas where change had occurred to water, urban build-up, green-space and bare lands, and four where no change had occurred. The class representing new urban build up was isolated from everything else and overlaid over the original 2005 image to reference what had occurred since 1994 in terms of urban development on the outskirts of the amalgamated City of Toronto (Fig. 5.8). From 1994 to 2005, approximately 164.07 km^2 of land has been disturbed (some excavated, some presently with urban build-up) which demonstrates that urban sprawl that has taken place during this time period. Much of the urban development occurred in municipalities surrounding the City of Toronto with a very small portion within the inner suburbs of Toronto (seen as dots).

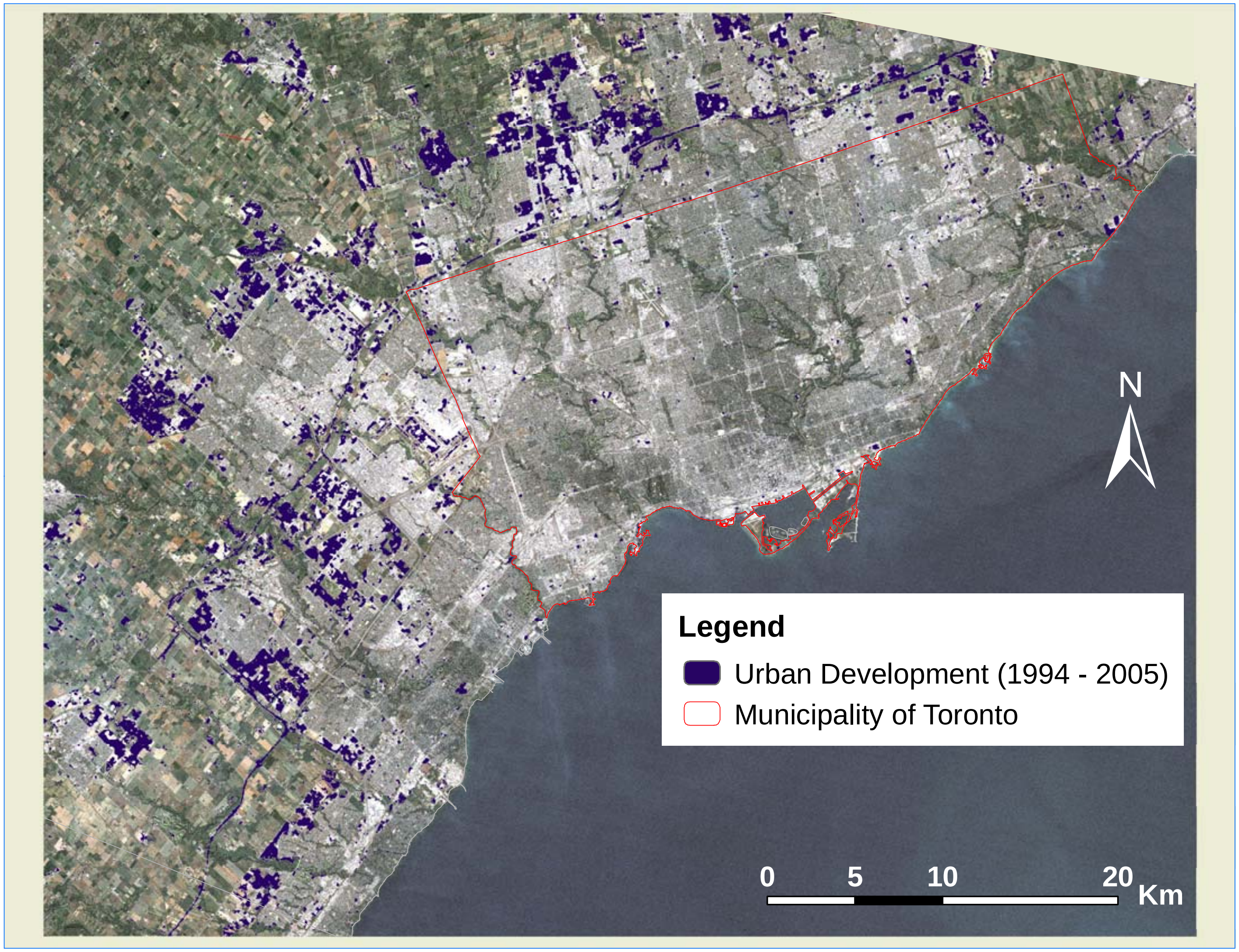

Figure 5.8 – Urban development that occurred between 1994 and 2005.

Methodology & Results (Part B)

5.4 Aerial ortho-photos, texture mosaics

This next section outlines the steps taken to obtain regressions between neighbourhood indicators and texture statistics. In order to capture sparse urban features which cannot be detected well at low resolution (Alhaddad et al., 2007) high resolution data in aerial photos was used. A more complete investigation would have used the same geographical area portion of the Toronto CMA used in Part A of the methodology to examine the spatial differences of the suburbs to the City of Toronto. However, ortho-photos making up an area that size would involve large data storage needs and lengthy data processing times. The City of Toronto provides an area extent where storage space and data processing time was just appropriate for the time provided to complete this research. A more inclusive version of this research would encompass a longer research period.

5.4.1 Ortho-photo mosaicing and texture analysis

Ortho-photos of the City of Toronto for 1997 and 2005 were collected with 187 images from 1997 and 640 images from 2005. Eight ortho-photos from 1997 were corrupted and were not used. The spatial resolution for 1997 was 0.5m and 0.2m for 2005. All images were mosaiced into one image for each year. The mosaic process involved the Mosaic tool in ERDAS Imagine software. Both mosaiced images were then transformed using a texture measure to display contrast texture across the city. The Contrast measure was chosen as the texture measure of choice after transforming mosaiced subsets for each year and looking for visual differences between urban build up features and all other features (Fig. 5.9 & 5.10). Contrast is the opposite of Homogeneity. Contrast measures the local variation in the image; it is high when the local window has high heterogeneity (PCI Inc., 2003).

As mentioned earlier, the size of the window used in a texture analysis depends on the resolution of the image. For this exercise the best window size was found through trials with varying grey levels. For 1997, the parameters for the contrast texture were a 15x15 window with 16 grey levels. For 2005, the parameters were a 21x21 window with 21 grey levels. These parameters best displayed differences between urban features and everything else. The urban developed areas in Figures 5.9 and 5.10 are displayed in high contrast values. This is because much of the dense urban areas incorporate many different land features and are not homogeneous in relation to less dense urban developed areas that have green-space and open-land. For example, residential areas are made up of concrete, trees and grass, as well as asphalt and materials that make up housing units and buildings.

The texture results may have been affected by the absence of adjacent ortho-photos for 1997. The contrast results could have benefited from the information in the unused ortho-photos mentioned. Also, cut lines between each photo were not smoothed out (an option that was not realized). By not smoothing out the cut lines, radiometric differences between photos were not reduced (PCI Geomatica, 2008). This may have allowed large variances of radiometric values to be included in the mean contrast calculation. Thus, it can be assumed that the mosaic may not represent accurate contrast values for 1997.

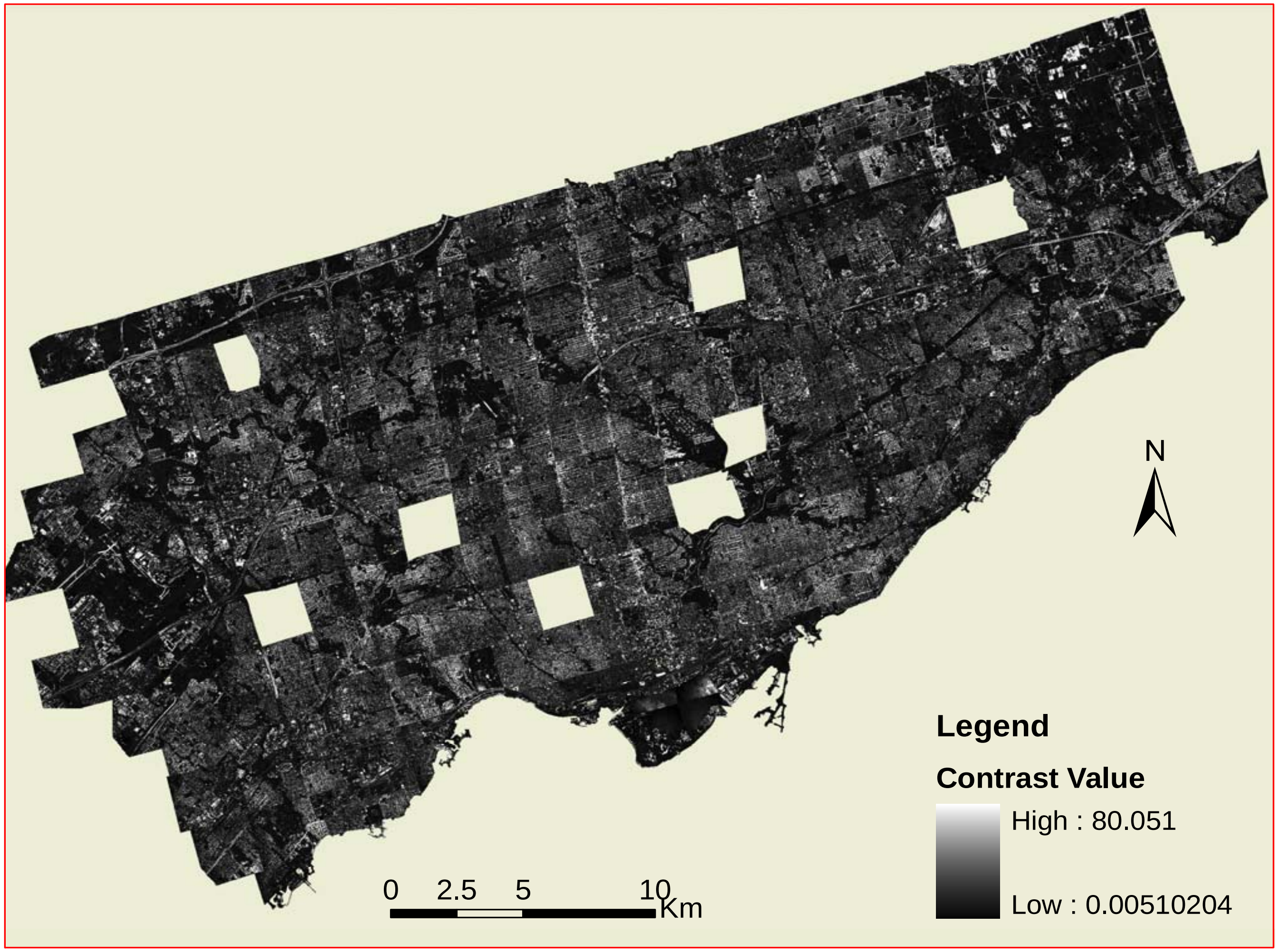

Figure 5.9 – Contrast texture for a mosaic of ortho-photos for the City of Toronto (1997)
*Produced by Sergio Barrios Jr. under License with the Ontario Ministry of Natural Resources © Queens Printer for Ontario, 20** 1997 and 2005*

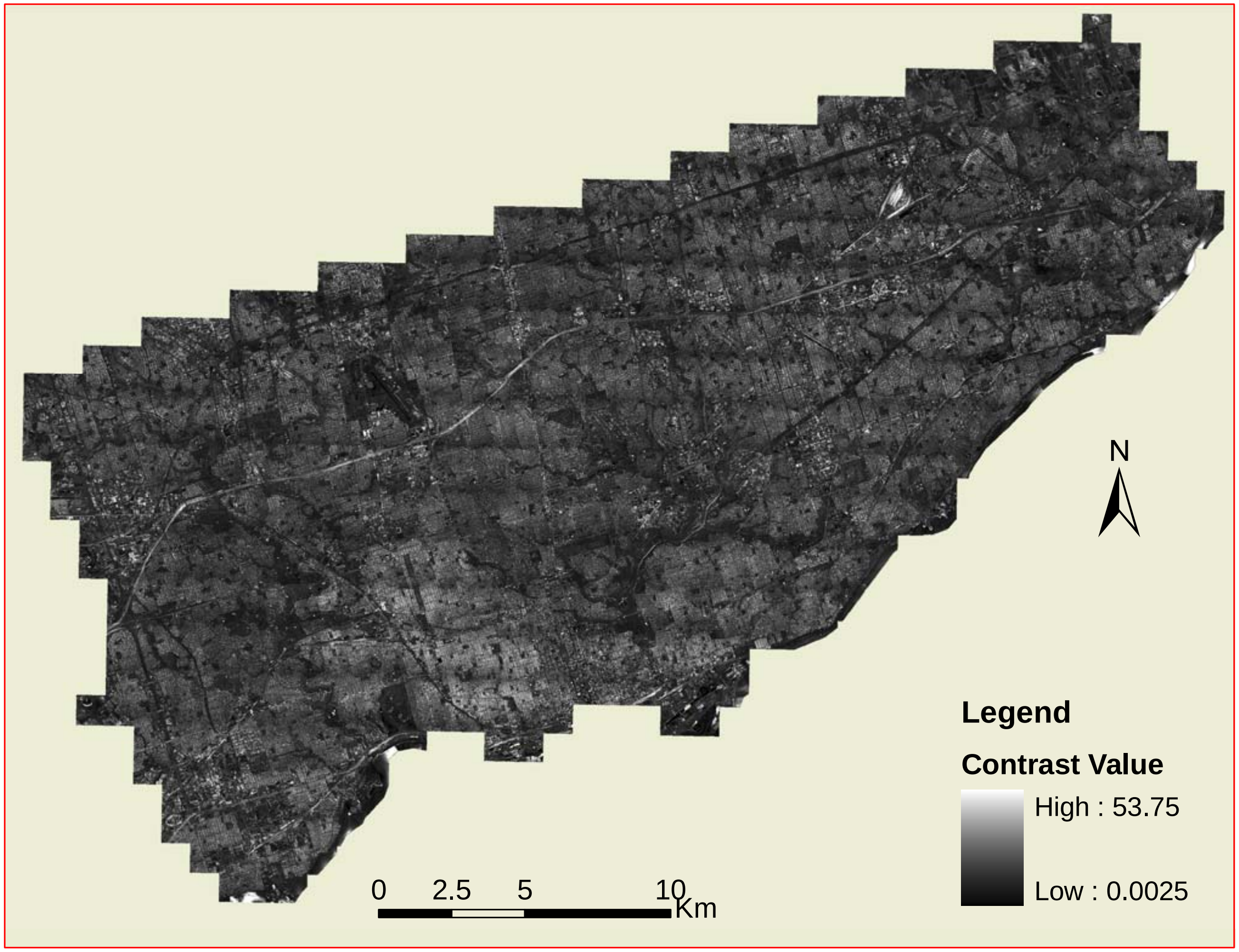

Figure 5.10 – Contrast texture for a mosaic of ortho-photos for the City of Toronto (2005)
Produced by Sergio Barrios Jr. under License with the Ontario Ministry of Natural
Resources © Queens Printer for Ontario, 20** 1997 and 2005

5.4.2 Zonal Statistics

The mosaic texture images had statistics calculated for them and attributed to polygons using the *Zonal Statistics to Table* tool in ESRI's ArcMap. The polygons were 1996 enumeration areas (EA) and 2006 dissemination areas (DA) for the City of Toronto. Both shapefiles were projected according to the mosaic images. The mean contrast was calculated for each EA and DA within new tables containing the same information found in the attribute tables of the shapefiles. Figure 5.11 shows the mean contrast textures for the City of Toronto for 1996 and 2006 shapefiles.

The difference in mean contrast between 1997 and 2005 is slight. Aside from the values becoming higher in 2005, the patterns are similar with darker areas of homogeneity found on the edges of the city and lighter areas of increasing heterogeneity found towards the centre. In 2005, however, increasing pockets of heterogeneity have formed as indicated by light grey to white areas (8.23 to 17.22).

It is interesting to note that the priority neighbourhoods seen outlined in red (Fig. 5.11) are mostly found in areas with relatively high mean contrast values (>8.23) for 2006. The mean was calculated for the contrast values for the three study area extents within the City of Toronto: thirteen priority neighbourhoods, all neighbourhoods in the City of Toronto excluding the thirteen priority neighbourhoods (hereon referred to as PNs Excluded), and all neighbourhoods (City of Toronto) (Table 5.4).

Priority neighbourhoods were found characteristic of lower mean contrast levels (lower heterogeneity) while PNs Excluded on average had higher heterogeneity out of the three for 1996 and 2006.

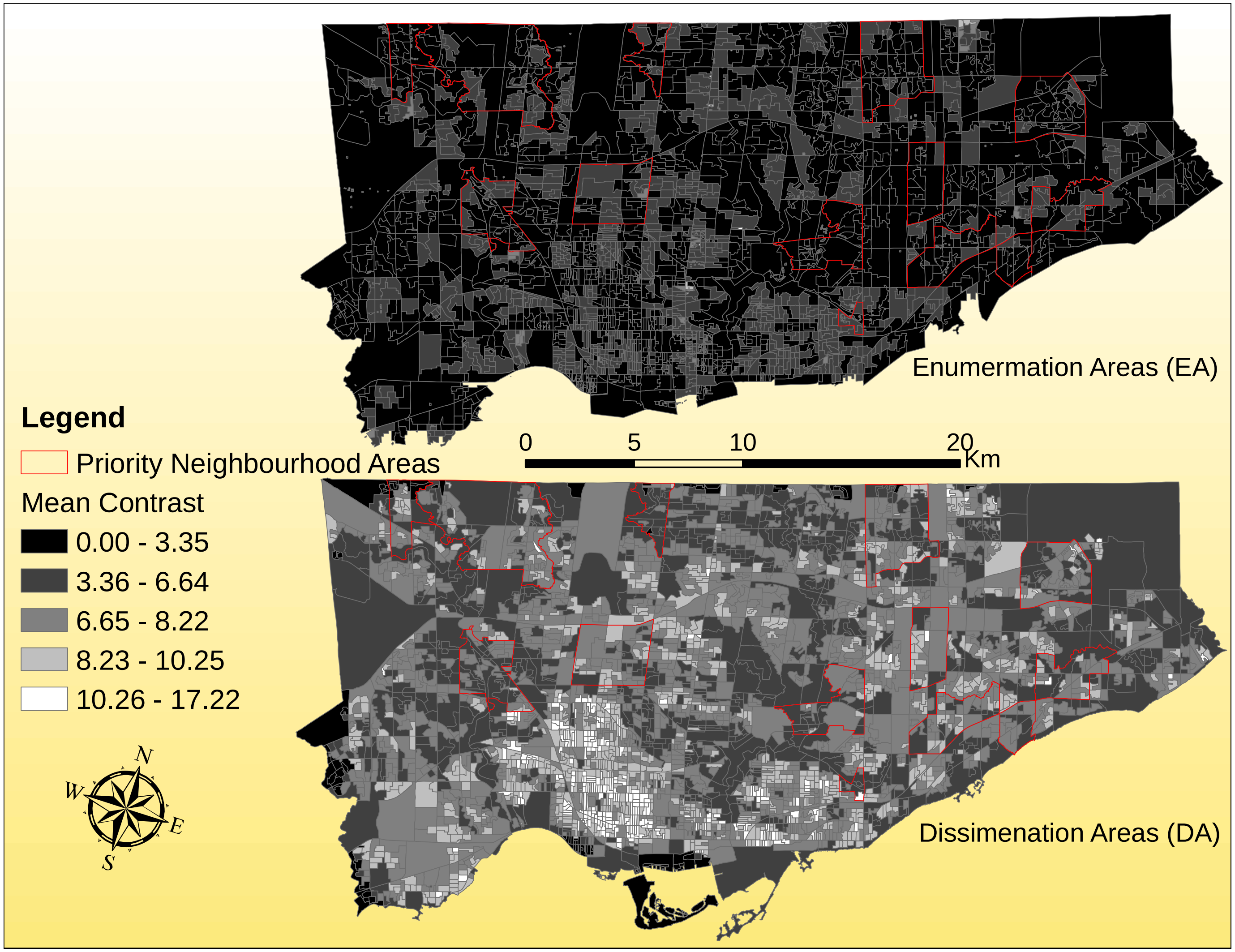

Figure 5.11 – Mean Contrast Texture for Enumeration Areas (1996) and Dissemination Areas (2006) for the City of Toronto

As contrast texture dictates, one wandering in areas of PNs Excluded would on average encounter surroundings that are more mixed in different surface materials than in the priority neighbourhoods.

Table 5.4 Mean mean contrast for the three study area extents of Toronto for 1996 and 2006

Year	Study Area Extent	Mean Mean Contrast
1996	Priority Neighbourhoods	2.7
	PNs Excluded	3.33
	Toronto	3.21
2006	Priority Neighbourhoods	6.96
	PNs Excluded	7.73
	Toronto	7.61

5.5 Census data

Census data for 1996 and 2006 were collected from CHASS (Computing in the Humanties and Social Science), a University of Toronto source for census data. The variables collected were those described as neighbourhood vitality indicators as stated in the report *Measuring Neighbourhood Vitality*, written for the City of Toronto by GHK Consulting (Dobilas and Battye, 2005). This report examined previous reports and practices concerning neighbourhood vitality studies so as to develop and apply the best indicators for Toronto. Its purpose was to assist in creating a tool that would identify the necessary attributes in order to define neighbourhood vitality so that decisions would be informed and priorities set in regenerating Toronto neighbourhoods. The report also reviews indicator systems at the neighbourhood level from international, national, and local practices, and finally sets a list of domains, each with indicators chosen to examine Toronto's neighbourhoods.

The report by Dobilas and Battye (2005) defines an indicator as “statistics or measures that provide evidence of conditions or problems”. The indicators are found as either qualitative or quantitative from varying data collection systems (e.g. customized vs. standard administrative). Indicators can also be regarded as neighbourhood asset (positive contribution) or drawback. An example is the percentage of dwellings needing major repair: high percentage possibly a deprivation and low percentage as an asset.

The list of indicators in the report was narrowed down from a longer list using careful consideration; criteria were drawn from previous work that undertook similar synthesis. Every chosen indicator warranted logic and justification.

Among the domains of indicators were Safety, Economy, Education, Urban Fabric, Health, and Demographics. Unfortunately a number of domains were inaccessible as they were not from census data, the only source of social/demographic data used in this report. The domains that were accessible were Economy, Education, Urban Fabric, and Demographics.

A variable selection problem arose because of the dissimilarity between the variable types available for the 1996 census and 2006 census. The aim of this part of the research was to not only identify relationships between remote sensing (texture) and GIS (census data), but a time series analysis of texture correlations with a specific set of social variables for a trio of neighbourhood study area extents. However, the 2006 census variable list is lengthier and more detailed than the 1996 census variable list. Some of the variable names did not match from 1996 to 2006 (renamed a little differently). Thus, a list was comprised from out of the domains available and matched from both census years. Table 5.5 contains the variable list mentioned as well as variable names that were assigned, their descriptions, strengths and weaknesses.

As a last note, missing or unavailable census data values for variables such as income or population were not included. Only those areal units which had values were used in the statistical analysis. This is but one method of handling census data, others estimate these missing values. One such example may take the weighted influence of values of neighbouring units depending on the size of the unit or the distance between centroids.

The indicator variables were then joined to either 1996 EA shapefile or the 2006 DA shapefile, depending on census year, and the two tables were then exported to SPSS (Statistical Package for the Social Sciences) for correlation analysis. Table 5.6 and 5.7 display the correlation values of each of the 53 variables to the mean contrast values for the three study area extents for 1996 and 2006.

Table 5.5 Selected List of Neighbourhood vitality indicators

Domain	Assigned Variable Name	Description	Strengths	Weaknesses
Demographics	Nonmovers	Nonmover status 1 year ago	Show stability of neighbourhood	Not clear who is moving
	Movers	Mover status 1 year ago	""	""
	Nonmigrants	Nonmirgrant status year ago	""	""
	Migrants	Migrant status 1 year ago	""	""
	IntnlMigrants	Internal migrant status 1 year ago	""	""
	ExtMigrants	External migrant status 1 year ago	""	""
	Nonmovers5	Nonmover status 5 year ago	""	""
	Movers5	Mover status 5 year ago	""	""
	Nonmigrants5	Nonmigrant status 5 year ago	""	""
	Migrants5	Migrant status 5 year ago	""	""
	IntnlMigrants5	Internal migrant status 5 year ago	""	""
	ExtMigrants5	External migrant status 5 year ago	""	""
	LoneParMale	Total Lone Male Parent	Useful predictor; children here have less good outcomes	Risk of stigma
	LoneParFem	Total Lone Female Parent	""	""
	15-19	% age 15-19	Useful predictor; begin at working age	
	20s	% age 20s	""	""
	30s	% age 30s	""	""
	40s	% age 40s	""	""
	50s	% age 50s	""	""
	60s	% age 60s	""	""
	70s	% age 70s	""	""
	80s+	% age 80s and older	""	""

	PerRecImm	Percent Recent Immigration (of total population)	Neighbourhood stability change	
	PerNonOffLang	Percent No knowledge of official languages (of total population)	Predictor for citizenship and access to services	
	PerNonHomLang	Percent Non-official home language (of total population)	Indicates access to services; linguistic changes - useful for planning?	Not show command of official languages
Urban Fabric	AutoDriver	Mode of Transport Automobile - Driver	Shows use of infrastructure - measure effects of transport plans	Sensitive to issues of access
	AutoPassen	Mode of Transport Automobile - Passenger	""	
	PublicTrans	Mode of Transport Automobile - Passenger	""	
	Walked	Mode of Transport Walked	""	
	Bicycle	Mode of Transport Bicycle	""	
	Motorcycle	Mode of Transport Motorcycle	""	
	Taxicab	Mode of Transport Taxicab	""	
	Other	Mode of Transport Other	""	
	Popden	Population Density (/km^2)	Good planning indicator	No clear relationship between density and overcrowding
	MajorRep	Number of Occupied Private Dwellings Requiring Major Repairs	Shows miss-match between supply of housing and demand to live in area	
	AvGrossRent	Average Gross Rent	Shows strength of demand and barriers to entry	every 5 years presents wide trends
	Owned	Tenure - Number of Owners	Shows mix of housing and tenures; and stability	
	Rented	Tenure - Number of Renters	""	
	GrRent30+	Gross Rent Spending: 30% or	Shows vulnerability and	every 5 years presents

		more of household income on shelter costs	housing lose risk	wide trends
	OwnMajPay30+	Owners Major Payment Spending: 30% or more of household income on shelter costs	Shows vulnerability	
Education	TEduNone	Total population with no education	Good predictor of later outcomes (and of children)	
	AttainNoGr9	Attainment of less than Gr.9 education, Pop. 15+	Good predictor of later outcomes	
	EduH	Attainment of highschool education	""	
	EduAppreTrd	Education - Apprentice/Trades	""	
	EduUni	Education - University	""	
	UniNoBach	Education - University, No Degree	""	
	UniYes	Education - University, Degree	""	
	UniBach+	Education - University, Bachelor Degree and more	""	
Economy	W_EI	Population with Employment Income	Proxy for measure of poverty	trends may be to changes in policy of criteria
	Emp_Rate	Employment Rate	Central factor in preventing social exclusion	
	Ump_Rate	Unemployment Rate	""	
	Avhhldinc	Average Household income	Proxy for median hhld income; key measure of social exclusion	
	AVINC	Average income	Added variable; may give some insight to distribution of wealth	Influence of extreme incomes

* Note: Strengths and weaknesses from Doblias and Battye (2005).

5.6 Stepwise Regression Analysis

Regressions show the direction and magnitude explanatory variables have on the dependent variable through coefficients. So, for purposes of this research, stepwise regression analysis was attempted with the mean texture contrast as the dependent variable. Six regressions were run: one for the entire City of Toronto, one for the Priority Neighbourhoods, and one for PNs Excluded for both 1996 and 2006.

Stepwise regression first starts with an empty model and then begins to add variables that are statistically significant (produces a significant F change). Each time a new variable is added to the model, the significance of variable already in the model is re-examined. In other words, each step tests the significance of each variable currently in the model and then the variable with the highest p-value (e.g. $p > 0.05$) is removed. The model is then refitted without that variable before proceeding to the next step (Murnan et al., 2004). Stepwise regression is a widely accepted technique in that it is able to consider more (relevant models) than other techniques (e.g. forward or backward regression). At the same time, stepwise is much quicker when the number of explanatory variables in the model is large (Larsen, 2008). In preparation for regression, homoscedasticity and collinearity were addressed.

5.6.1 Homoscedasticity

Homoscedasticity assumes that the variance found in the dependent variable is equal to those found in the rest of the data (independent variables) (Princeton University, 2007). The dataset used here did not have equal variances (heteroscedasticity) as seen through descriptive statistics. Performing a regression with heteroscedastic data will estimate regression coefficients that are still unbiased but it is less efficient than those calculated from data that have been transformed (Ghilagaber, 2004). By using transformations, such as logarithmic or square-root for positive skewness, and powers for negative skewness, heteroscedasticity can be reduced to a reasonable level. Descriptive statistics were run in SPSS to examine the extent of skewness. Those with significantly positive skewness were transformed by square-root (as logarithmic transformations were not of assistance) and those with significantly negative skewness were transformed by powers. The option of removing outliers from the dataset was considered but not performed as values were spatially attached to EA and DA polygons which are spatially related. The analysis here requires the dataset to be as spatially whole as possible.

5.6.2 Collinearity

Collinearity, or multicollinearity, is defined in econometrics as the presence of close association (linear or near linear relationships) among explanatory (independent) variables in a linear regression (Silvey, 1969). The purpose of regression analysis is "to estimate parameters of a dependency, not an interdependency relationship", (Farrar and Glauber, 1967). Interdependency causes problems in regression because the high associations may have certain variables appear significant to the analysis when in truth they are only significant because of another variable included in the regression (Rawlings et al., 1998). This appears in inflated estimated coefficients with an incorrect sign at times. This can lead to unreasonable predicted values. Regressions, such as stepwise, may fail to recognize vital variable combinations (Hocking, 1983). It is important to examine collinearity before performing a regression so that issues caused by interdependency as mentioned above are reduced. After focusing on homoscedasticity, collinearity within the independent variables was observed through a correlation matrix and determinant value in SPSS through factor analysis. Li and Weng (2007) used a similar method to isolate high correlations between variables as mentioned before. As with factor analysis, collinearity can cause problems in regression because it becomes difficult to determine the contribution of variables that are highly correlated (Field, 2005). Li and Weng went on to use create indices based on aspects of quality of life (Economic, Environmental, and Crowdedness) from the factor loadings in principal components having eigenvalues greater than 1. This step was not performed as domains of neighbourhood indicators were already created.

The following steps for observing collinearity are taken from Chapter 15 in Field (2005). A correlation matrix is produced showing the interrelationships between every indicator variable. In the matrix $R > 0.8$ are considered a high association as a rule of thumb (R representing the correlation coefficient). One method of reducing the high associations is the removal of variables with said R value. A high number of high associations were found between pairs of variables. The question then became: *which of these two variables accounts for more in the other and would be best suited to then stay in the dataset?* By removing the variable of the pair that is least accounted for in the other removes the chance that the removed variable would be included in the regression model because of its high association. This approach involves some decisions to be some what biased on the part of the researcher by removing variables to reduce collinearity when the amount of association between a pair of variables is significantly high. The determinant statistic indicates the presence of collinearity overall. A value of less than 0.00001 indicates that collinearity is still present. This standard could not be sufficiently met in an attempt to isolate one variable list for all three study area extents. Thus, a sub sample of only four variables was selected from the fifty-one variables where collinearity would be low.

The significance level used was 0.05 which means there is a 5 percent chance that a sampling error was made. It is a commonly accepted level. A lower significance level would lower the chances of committing a sampling error but the analysis becomes less precise. A higher significance level thus, increases the chance of committing a sampling error. A significance level of 0.05 in a unique, experimental analysis of this type seems appropriate (McGrew, Jr. and Monroe, 2000).

5.6.3 Stepwise Regression Results

The four variables used were average household income (Avhhldinc), attainment of high school education (EduHS), percentage of recent immigration (PerRecImm), and population density (Popden). These four were chosen for regression so as to have some indication of the value of contrast texture in social science investigations. Each one was taken from one of the four neighbourhood indicator domains. Correlations between the variables did not exceed ± 0.578 indicating low collinearity. The results of the stepwise regression models are found in Table 5.6.

The R square measures the amount of variance found in the dependent variable (i.e. mean contrast) that can be explained by the explanatory variables entered in a model. The adjusted R square measures the proportion of the variation in the dependent variable accounted for by the explanatory variables. Adjusted R square is generally considered as more accurate in measuring the goodness-of-fit than R square. The reason is adjusted R square allows for the degrees of freedom associated with the sums of the squares. Also, when the adjusted R square is significantly lower than R square this normally means that some explanatory variable(s) are missing, thus not fully measuring the variation in the dependent variable. The adjusted R square can be used to compare model results which have a differing number of observations or independent variables (Natural Resources Canada, 2005).

All variances explained by the indicators are very small with the highest adjusted R square being 0.089 in the 2006 PNs Excluded model. Each model, however, had very similar R square and adjusted R square values for each indicator. Thus, the variation in the mean contrast may have been fully measured by the indicators that were included each in model. In the 1996 Priority Neighbourhoods model, for example, PerRecImm was the only indicator included which means it was the only indicator of the four chosen that was statistically significant, i.e., PerRecImm's regression coefficient (-0.14) is significant at the 0.05 level. The other three indicators were not significant ($p>0.05$) and so not included. The regression coefficient indicates the direction and strength of the relation between an explanatory variable and the dependent variable. For PerRecImm, a negative sign indicates an inverse relationship, i.e. were the percent of recent immigration for priority neighbourhoods to decrease by 0.14 the value of mean contrast texture (or heterogeneity) would increase by 1. In other words, where heterogeneity would have been high the percent of recent immigration would be lower for priority neighbourhoods in 1996. The model's R square (0.02) and adjusted R square (0.019) are not significantly different. Therefore, PerRecImm measures the variation of the mean contrast well enough and no other indicators were needed.

A notable point is that two of the three 2006 year models had all four chosen indicators included in their models, whereas one 1996 model had two indicators (a maximum among 1996 models). In addition, the combinations of variables that are able to fully explain the variance are different from 1996 to 2006 for corresponding models (e.g. priority neighbourhoods 1996 have PerRecImm; priority neighbourhoods 2006 have EduHS and Popden).

Table 5.6 – Stepwise Regression Model Summaries.

*Note: Dependent variable - mean contrast texture.

Model				
		Coefficient	**R^2**	**Adjusted R^2**
1996 Priority Neighbourhoods	Constant	-0.017		
	PerRecImm	-0.14 (*)	0.02	0.019
1996 PNs Excluded	Constant	6.11E-05		
	EduHS	-0.063(*)	0.004	0.004
1996 All of Toronto	Constant	-0.005		
	PerRecImm	-0.085(*)	0.009	0.008
	EduHS	-0.066(*)	0.013	0.012
2006 Priority Neighbourhoods	Constant	0.696		
	EduHS	-0.180(*)	0.048	0.046
	Popden	-0.120(*)	0.061	0.058
2006 PNs Excluded	Constant	0.008		
	EduHS	-0.225(*)	0.044	0.043
	Popden	0.204(*)	0.077	0.076
	PerRecImm	-0.120(*)	0.086	0.085
	Avhhldinc	-0.068(*)	0.09	0.089
2006 All of Toronto	Constant	0.005		
	EduHS	-0.229(*)	0.052	0.052
	Popden	0.154(*)	0.069	0.068
	PerRecImm	-0.128(*)	0.079	0.078
	Avhhldinc	-0.061(*)	0.083	0.082

(*) Significant at 0.05 level.

Chapter 6: Discussion

6.1 Research results

6.1.1 Change Detection – Urban Sprawl for Toronto

Part A of the Methodology and Results identified urban build-up between 1994 and 2005 outside the City of Toronto as seen in Figure 5.10. In this 11 year time-frame approximately 164.07 km^2 of land had changed into urban build-up within the portion of the Toronto CMA examined in this report. This change in land due to urban development represents 54% of the total land change that had occurred. A very small portion of this change occurred within the inner suburbs of the Toronto, an amount that does not seem reasonable to blame for the reported increased poverty/inequality found today. However, the greater portion of sprawl in the surrounding municipalities should be examined in relation to the City of Toronto as to how much influence urban development in these areas may have on poverty today.

6.1.3 Stepwise Regression

One possible explanation why collinearity could not be reduced is that each study area extent may have its own unique dataset that does not exhibit collinearity and that no one data set is suitable for all three. However, the process of finding unique datasets for each study area extent would be a great task using the method outlined earlier (especially if the initial dataset is lengthier to begin with). At this point it is obvious that this approach in handling collinearity through a correlation matrix for a large data set is not appropriate. Li and Weng used a correlation matrix for such a purpose but for a data set of only a dozen variables where collinearity may have been low to insignificant, compared to the fifty-one variables used here where collinearity would have a better chance of being present. It seems that the variables collected from the census are too highly associated with each other and that a more efficient, accurate method of removing collinearity from a large variable data set for regression analysis is called for.

The regression models were fairly weak with 0.089 as the greatest variance explained by any of the models. The combinations of explanatory variables were also different between corresponding models for both years. This may implicate that a shift in what may account for variances in heterogeneity had taken place between 1996/97 to 2005/06.

As mentioned before, the 1997 mosaic lacked several photos and had visible cut lines where photos joined which may have had an influence on the texture analysis as well as the regression results in the 1996 models.

Chapter 7: Conclusion

Census data and remotely sensed imagery are essential sources of data in urban analyses. Remote sensing data record the physical properties of the environment and provide large variable descriptions in order to reveal significant regression results with texture as the dependent variable. Census data offer a variety of demographic and socio-economic information used in inequality research (Frey, 2001), urban planning and management. GIS has advanced to where spatial analysis of remotely sensed data and other sources of spatial data can be performed in an effective environment (Donny et al., 2001). Literature on relating remotely sensed data to census data for the purpose of using remotely sensed data as an indirect indicator of social area vitality is not abundant (see Li and Weng 2007 and Jensen and Gatrell 2005). This research attempted to do so using contrast texture and neighbourhood vitality indicators created for the City of Toronto.

There were a number of factors that would have had an influence on the outcome of the analysis. The fact that contrast texture was averaged for each census polygon means that any large variance in texture would have been simplified through aggregation in each polygon. Texture information may have been lost in doing so and any conversion between raster to vector, or vice versa, can have that effect. Finding desirable aggregation units is important to reduce the loss of information. Also, the parameters used to produce the contrast values from the ortho-photo mosaics were user influenced. Since the settings of these parameters (grey levels, size of window) are dependent on the remote sensing platform too, similar research is bound to have varying results. Another issue is the pre-analysis data preparation of the census variables. Other research may opt to use different tactics such as filling in suppressed census data instead of leaving those areas out of the analysis. In a larger study, this approach may have improved the results; however, there are several things that were excluded due to this reason and described in the next chapter.

Texture analysis in change detection was first performed to show that suburban sprawl had taken place. Suburban sprawl was once to blame for some of the neighbourhood poverty/inequality that had taken place in the City of Toronto almost 5 decades ago. As priorities are now focused on re-building the social/physical infrastructure of neighbourhoods, suburbs may no longer be argued as a poverty mechanism. Over time this trend may have dissipated with changing economics, reduced crime, neighbourhood cultural blossoming, and on going gentrification within the city. However, it is possible that the degree to which suburban sprawl affects poverty can fluctuate over time in the City of Toronto. That is a subject for another research.

Perhaps the greatest finding this research may have uncovered is an area's heterogeneity may influence, or be influenced, by variables that change over time. None of the corresponding models between 1996 and 2006 had the same combinations of significant indicators. However, had collinearity been reduced sufficiently between the 53 indicators more impressive results may, or may not, have been produced.

The magnitude that the missing photos and un-smoothened cut lines had on the 1996 regression results remains to be seen. In speculation, the City of Toronto and its neighbourhoods may be under going a transition which may explain the weak regression models. Or perhaps there may be no significant relation between the spatial distribution of heterogeneity and these vitality indicators within the City of Toronto. Meaning it is possible this type of analysis is not meant for the City of Toronto in particular. Other cities may be similar to the City of Toronto in this regard, and others may in fact be more "organized" with respect to heterogeneity. Developing world cities may be quite different from developed cities like Toronto, from the cost of living to mobility which often reflects differing approaches to public policies. In addition, population projections show some of the largest cities in 2015 will be in the developing world while some large cities in the developed world may experience population declines (Bugliarello, 1999). Applying an analysis such as the one attempted here to locations of differing policy making, larger populations, and customs may be very difficult.

Chapter 8: Recommendations

8.1 Recommend Approaches

8.1.1 Recommended Regression Approach: Principal Component Regression

Multicollinearities are usually, but not always, indicated by large correlations between pairs of variables. Various methods exist for choosing a subset that does not contain multicollinearities (Draper and Smith, 1998; Hocking, 1976; Miller, 1984, 1990 as seen in Jolliffe, 2002). Some of these methods are based on principal components with the best-known approach generally known as Principal Component (PC) regression. It is a method that combines linear regression with principal component analysis (Draper and Smith, 1981). It begins by using the principal components of the independent variables in place of the independent variables. It gathers highly correlated independent variables into a principal component with all principal components independent of each other. Essentially, this method transforms a set of correlated variables into a set of uncorrelated principal components (Liu et al., 2003). There are different ways on how to select the number of PCs for regression, each with their own setbacks as described in chapter 8 in Jolliffe (2002). What PC regression can do that ordinary least squares regression cannot is to explicitly indicate whether the removal of multicollinearity is problematic, i.e., if removing a substantial number of PCs will solve instability in the regression coefficients.

A detailed step-by-step method in SPSS can be looked up in Liu et al. (2003). Here, the regression method is completed through six steps. They discuss the procedure by stating that not only can collinearity be diagnosed and overcome using PC regression, but any variable's relation to the dependent variable that conflicts with any *a priori* knowledge can be corrected as well while preserving the original information.

8.1.2 Other Recommendations

There were other approaches that were not apart of this report but may be of interest to future investigations. Aside from involving more data and applying PC regression, another approach would be to perform a more detailed time series analysis involving multiple mosaics, some from before 1997. Doing so may capture a trend or shifts in heterogeneity-indicator(s) relations.

Another approach would be to include the CMA in the texture-to-social variable analysis. This would give a wider scope as to where inequality has spread and where it has not. The one difficulty in doing so is the amount of data storage and computer processing power required when using aerial photos. The operational issue of making any corrections/adjustments made to the texture analysis can be costly in processing time. Finally, utilizing the option to smooth out cut lines in the mosaic may produce more reliable texture results.

A final recommendation is, if research is able to produce significant relations between heterogeneity (or perhaps another type of remotely sensed data) and census variables, to access neighbourhoods individually instead of a generalizing them. Many neighbourhoods may be similar in a number of regards, and generalizing may prove as time and cost efficient when accessing their needs, however, they each may have differences. Therefore, each neighbourhood may require different services and at varying scales.

REFERENCES

Alhaddad, B.I., Burns, M.C., and Cladera, J.R. (2007). Texture Analysis for Correcting and Detecting Clasification, Structures in Urban Land Uses: “Metropolitan area case study – Spain”. *Urban Remote Sensing Joint Event*. Retrieved July 15, 2007 from http://www-cpsv.upc.es/documents/URS2007_II.pdf.

Anderson, I. (2008). Feature Extraction and Classification: Solutions Paper. Retrieved August 1, 2008 from http://www.erdas.com/files/whitepapers/Feature_Extrt_Class.pdf

Armenakis, C., Ledu, F., Cyr, I., Savopol, F., and Cavayas, F. (2003). A comparative analysis of scanned maps and imagery for mapping applications. *Journal of Photogrammetry & Remote Sensing*, 57, 304-314.

Azimi-Sadjadi, M. R., Ghaloum, S., and Zoughi, R. (1993) Terrain classification in SAR images using principal components analysis and neural networks. *T ransactions on Geoscience and Remote Sensing*, 31, 511–515.

Benediktsson, J. A., and Sveinsson, J. R. (1997). Feature extraction for multisource data classification with artificial neural networks. *International Journal of Remote Sensing*, 8, 727–740.

Bruzzone, L, and Bovolo, F. (2007). Unsupervised Change Detection in Multi-Temporal SAR Images. In Chen, C.H. (Eds.) *Signal and Image Processing for Remote Sensing*. Boca Raton, FL.: CRC Press, Taylor & Francis Group.

Bugliarello, George. (1999). Megacities and the Developing World. *The Bridge*, 3. Retrieved September 12, 2008 from http://www.nae.edu/nae/bridgecom.nsf/weblinks/NAEW-4NHMPU?OpenDocument

CANSIM (2007). Canadian Census Analyser @ CHASS: Canadian Census Profile (Enumeration Area/Dissemination Area level). Using CHASS at the University of Toronto.

Centre For Urban and Community Studies, Univeristy of Toronto. (2007). The Three Cities within Toronto: “a city of disparities”. Retrieved September 11, 2008. from http://www.fairlawnavenueunited.ca/social_justice_city_of_disparities.pdf

Chen, K. (1998). Correlations between Census Dwelling Data and Remotely Sensed Data. In *The 10th Annual Colloquium of the Spatial Information Research Centre, 16-19 November 1998* (pp. 1-9). University of Otago, Dunedin, New Zealand.

Chen, K. (2002). An Approach to linking remotely sensed data and areal census data. *International Journal of Remote Sensing*, 23(1), 37-48.

City of Toronto (2005). City of Toronto Priority Neighbourhoods (boundary file: Priority Neighbourhooods). From Social Policy Analysis & Research, City of Toronto.

City of Toronto. (2008). Toronto's Agenda for Prospertiy. Retrieved on September 11, 2007, from http://www.toronto.ca/prosperity/index.htm#prosperity

Cheng, J., and Masser, I. (2003). Urban growth pattern modeling: a case study of Wuhan City, PR China. *Landscape and Urban Planning*, 62, 199-217.

Congalton, R.G. (1991). A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing of Enviroment*, 37(1), 35-46.

Conners, R. W., and Harlow, C. A. (1980). A theoretical comparison of texture algorithms. *Transactions On Pattern Analysis and Machine Intelligence*, PAMI-2, 204–222.

Dark, S.J., and Bram, D. (2007). The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography*, 31(5), 471-479.

De Paul, O.V. (2007). Remote Sensing: New Applications for Urban Areas. *Proceedings of the IEEE*, 95(12), 2267-2268.

Dobilas, G., and Battye, F. (2005). *Measuring Neighbourhood Vitality*. Toronto, Ontario, Canada: GHK International (Canada) Ltd..

Draper, N.R., and Smith, H. (1981). Applied Regression Analysis (2nd ed.). New York: John Wiley & Sons Inc.

Farrar, D.E., and Glauber, R.R. (1967). Multicollinearity in Regression Analysis: The Problem Revisited. *The Review of Economics and Statistics*, 49(1), 92-107.

Field, A.P. (2005). *Discovering statistics using SPSS (2nd ed.)* London: Sage.

First Base Solutions (2005). City of Toronto Orthophotos [digital resource: raster]. Toronto, Ontario: First Base Solution. Retrieved June 19, 2008, from Ryerson University Library.

Forsythe, K.W. (2003). Monitoring "Megacity" Urban Development: Pansharpened Landsat 7 Results for Toronto, Ontario, Canada. Retrieved May 31, 2008, from http://www.geography.ryerson.ca/wayne/SA8922-W2008/Forsythe-URS2003-Regensburg.doc.

Forsythe, K.W. 2004. Pansharpened Landsat 7 Imagery For Improved Urban Area Classification. *Geomatica*, 58(1).

Forsythe, K.W., and Waters, N.M. (2006). The Utilization of Image Texture In Urban Change Detection. *Photogrammetrie • Fernerkundung • Geoinformation*, 4, 287-296.

Frey, H.W. (2001). Melting pot suburbs: A census 2000 study of suburban diversity. *Census 2000 Series*, June. Washington, D.C.: The Brookings Institution.

Ghilagaber, G. (2004). Another Look at Chow's Test for the Equality of Two Heteroscedastic Regression Models. *Quality & Quantity*, 38, 81-93.

Glennerster, H. (2000). US Poverty Studies and Poverty Measurement: The past twenty-five years. Retrieved on September 11, 2008, from http://sticerd.lse.ac.uk/dps/case/cp/CASEpaper42.pdf

Goliath. (2008). Municipal managed gentrification in South Parkdale, Toronto. *The Canadian Geographer*, Retrieved on September 11, 2008, from http://goliath.ecnext.com/coms2/gi_0199-2355533/Municipally-managed-gentrification-in-South.html

Heisz, A., and Mcleod, L., Statistics Canada. (2004). Low Income in Census Metropolitan Areas, 1980-2000. *Trends and Conditions in Census Metropolitan Areas*, Business and Labour Market Analysis Division. Retrieved May 30, 2008, from http://www.statcan.ca/english/research/89-613-MIE/2004001/89-613 MIE2004001.pdf

Hocking, R.R. (1983). Developments in Linear Regression Methodology: 1959-1982. *Technometrics*, 25(3), 219-230.

Hoffer, R.M. (1978). Biological and physical considerations in applying computer-aided analysis techniques to remote sensor data. In Swain, P.H., and Davis, S.M. (Eds.) *The quantitative approach*. U.S.A.: McGraw-Hill.

Hong, Z., Jiang, Q., Dong, H., Wang, S., and Li, J. (2005). An Improved FCM-Based Model for Urban Change Detection Using High-Resolution Remotely Sensed Images. *Environmental Informatics Archives*, 3, 352-359.

Hutchinson, C.F. (1982). Techniques for combining Landsat and ancillary data for digital Classification improvement. *Photogrammetric Engineering& Remote Sensing*, 48, 123-130.

Income Statistics Division, Statistics Canada (2007). Low Income Cut-offs for 2006 and Low Income Measures for 2005. Retrieved May 31, 2008, from http://www.statcan.ca/english/research/75F0002MIE/75F0002MIE2007004.pdf
.

Irons, J.R., and Petersen, G.W. (1981). Texture Transforms of Remote Sensing. *Remote Sensing of Environment*, 11, 359-370.

Jargowsky, P.A. (2002) Sprawl, Concentration of Poverty, and Urban Inequality. In

Squires, G.D. (Eds.) *Urban Sprawl: Causes, Consequences & Policy Responses*. Washington, D.C.: The Urban Institute Press.

Jensen, J.R. (1996). Introductory to digital image processing: A remote sensing perspective (2nd ed.). New Jersey, U.S.A: Prentice Hall.

Jensen, J.R. (2000). Remote Sensing of the Environment. Upper Saddle River, New Jersey: Prentice Hall.

Jensen, R.R., & Gatrell, J.D. (2005). Image Homogeneity and Urban Demographics: An Integrated Approach to Applied Geotechniques. In Jensen, R.R., Gatrell, J.D., & Mclean, D.D. (Eds.) *Geo-spatial Technologies in Urban Environments*. Springer Berlin Heidelberg.

Jolliffe, I.T. (2002). Principal Component Analysis (2nd ed.). New York: Springer-Verlag.

Kotler, P., Roberto, N., and Leisner, T. (2006). Alleviating Poverty: A Macro/Micro Marketing Perspective. *Journal of Macromarkting*, 26(2), 233-239.

Land Information Toronto. (1997). City of Toronto Orthophotos: Orthophotos [digital resource: raster]. Toronto, Ontario: Land Information Toronto. Retrieved June 19, 2008, from Ryerson University Library.

Li, G. and Weng, Q. (2007). Measuring the quality of life in city of Indianapolis by Integration of remote sensing and census data. *International Journal of Remote Sensing*, 28(2), 249-267.

Liu, R.X., Kuang, J., Gong, Q., and Hou, X.L. (2003). principal component regression analysis with spss. *Computer Methods and Programs in Biomedicine*, 71, 141-147.

Lo, C.P., and Faber, B.J. (1997). Integration of Landsat Thematic Mapper and Census Data for Quality of Life Assessment. *Remote Sensing of the Environment*, 62, 143-157.

Lorinc, J. (2001). The story of sprawl. *Toronto Life*. May, 131-135.

Low, Harvey. (2008). *A General Overview of Income Measures: The Pros and Pitfalls of Poverty Measures* [PowerPoint]. Social Policy Analysis & Research Section, Social Development, Finance and Administration Division, City of Toronto.

Lui, X., and Lathrop, JR., R.G. (2002). Urban change detection based on artificial neural network. *International Journal of Remote Sensing*, 23(12), 2513-2518.

Martin, D. (1989). Mapping Population Data from Zone Centroid Locations. *Transactions of the Institute of British Geographers*, 14(1), 90-97.

MacDonnell, S. (2007). Losing Ground: The Persistent Growth of Family Poverty In Canada's Largest City. *A Report of United Way of Greater Toronto.* Retrieved May 29, 2008, from http://www.unitedwaytoronto.com/whoWeHelp/reports/pdf/LosingGround-fullReport.pdf

McGrew, Jr., J.C., and Monroe, C.B. (2000). An Introduction to Statistical Problem Solving in Geography (2nd Ed.). U.S.A.: McGraw-Hill Companies, Inc.

Megaw, W.J., Hirt, M.S., and Muller, E.F. (1974). Air Pollution Over Toronto. *Water, Air, & Soil Pollution*, 3(4), 483-488.

Mkhabela, M.S., Mkhabela, M.S., and Mashinini, N.N. (2005). Early maize yield forecasting in the four agro-ecological regions of Swaziland using NDVI data derived from NOAA's-AVHRR. *Agricultural and Forest Meteorology*, 129, 1-9.

Murnan, J., Dake, J.A., and Price, J.H. (2004). Association of Selected Risk Factors in Child and Adolescent Firearm Mortality by State. *The Journal of School Health*, 74(8), 335-340.

Natural Resources Canada. (2005). Correlation Coefficient and Coefficient of Variation. Retrieved on September 11, 2008, from http://www.pfc.cfs.nrcan.gc.ca/profiles/wulder/mvstats/correl_coef_e.html

PCI Geomatica. (2008). Release Notes: Geomatica 10.0.3. Retrieved September 3, 2008, from http://www.pcigeomatics.com/services/support_center/updates/pdf/GeomaticaV10.0.3_Release_Notes.pdf

PCI Inc. 2003. PCI Help Menu. Richmond Hill, Ontario.

Pesaresi, M. (2000). Texture Analysis for Urban Pattern Recognition Using Fine-Resolution Panchromatic Satellite Imagery. *Geographical and Environmental Modelling*, 4(1), 43-63.

Pusissant, A., Hirsch, J., and Weber, C. (2005). The Utility of Texture Analysis to Improve Per-Pixel Classification for High to Very High Spatial Resolution Imagery. *International Journal of Remote Sensing*, 26(4), 733-745.

Rawlings, J.O., Pantula, S.G., and Dickey, D.A. (1998). Applied Regression Analysis. (2nd ed.). New York: Springer London.

Ridd, M.K., and Liu, J. (1998). A Comparison of Four Algorithms for Change Detection in an Urban Environment. *Remote Sensing of the Environment*, 63, 95-100.

Rouse, J.W., Hass, R.H., Schell, J.A., and Deering, D.W. (1974). Monitoring Vegetation Systems in the Great Plains with ERTS. *Proceedings*, Third Earth Resources Technology Satellite-1 Symposium, Greenbelt: NASA SP-351, 3010-317.

Sadler, G.J., and Barnsley, M.J. (1990). Use of population density to improve classification accuracy in remotely sensed images of urban areas. In *Proceedings of EGIS '90*. Amsterdam: EGIS Foundation.

Shaban, M.A., and Dikshit, O. (2001). Improvement of classification in urban areas by the use of texture features: the case study of Lucknow city, Uttar Pradesh. *International Journal of Remote Sensing*, 22(4), 565-593.

Silvey, S.D. (1969). Multicollinearity and Imprecise Estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 31(3), 539-552.

Singh, A. (1989). Review Article: Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing*, 10(6), 989-1003.

SPSS Inc. (2006). SPSS Help Menu. Chicago, Illinois.

Statistics Canada (1996). Census of Canada 1996 Digital Cartographic File (cartographic boundary file, gea_535a.exe). Using University of Toronto Data Library Service.
Statistics Canada, 2002. Geographic Units: Census Metropolitan Area (CMA) and Census Agglomeration (CA). Retrieved August 1, 2008 from http://www12.statcan.ca/english/census01/products/reference/dict/geo009.htm

Statistics Canada (2006). Census of Canada 2006 Digital Cartographic File (cartographic boundary file, COT_DA_2006.zip). Using Ryerson University Geospatial Map and Data Centre.

Statistics Canada, (2006). The Daily: Building Permits. Retrieved May 29, 2008, from http://www.statcan.ca/Daily/English/060306/d060306a.htm

Statistics Canada, Household Surveys Division (1998). Low income cut-offs. Retrieved July 29, 2008, from http://www.statcan.ca/english/freepub/13-551-XIB/13-551-XIB1998001.pdf.

Solomon, L. (2007). Toronto Sprawls. Canada: University of Toronto Press Incorporated.

Squires, G.D. (2002) Urban Sprawl and the Uneven Development of Metropolitan America. In *Squires, G.D.* (Eds.) *Urban Sprawl: Causes, Consequences & Policy Responses*. Washington, D.C.: The Urban Institute Press.

Story, M., and Congalton, R. (1986). Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing*, 52(3), 397-399.

Strong Neighbourhoods Task Force (2005). *Strong Neighbourhoods: A Call to Action*…2005. Retrieved from http://www.toronto.ca/demographics/sntf/sntf_report.pdf

Sun, H., Forsythe, W., and Waters, N. 2007. Modeling Urban Land Use Change and Urban Sprawl: Calgary, Alberta, Canada. *Networks and Spatial Economics*, 7, 353-376.

Toronto City Summit Alliance (2007). *Strong Neighbourhoods: Supporting The Call To Action*. Discussion paper for Toronto Summit 2007, 26-27 February 2007. Retrieved from the United Way Toronto Web site: http://www.unitedwaytoronto.com/whoWeHelp/reports/SNTF/TCSA-Long_Paper.pdf

United Way of Greater Toronto, The. (2004). Poverty by Postal Code: New United Way Report "Poverty by Postal Code" Documents Changing Toronto Neighbourhoods. Retrieved on August 28, 2008 from http://dawn.thot.net/poverty-report.html#1

United Way, The. (2008). What We do. Retrieved May 30, 2008, from http://www.uwgt.org/whatWeDo/whatWeDo.php

Van Kempen, E.T. (1994). The dual city and the poor: Social polarisation, social segregation and life chances. *Urban Studies*, 31(7), 995-1015.

Vogelmann, J.E., Sohl, T., and Howard, S.M. (1998). Regional Characterization of Land Cover Using Multiple Sources of Data. *Photogrammetric Engineering & Remote Sensing*, 64(1), 45-57.

Weng, Q. (2002). Land use change analysis in the Zhujian Delta of China using satellite remote sensing, GIS and stochastic modelling. *Journal of Environmental Management*, 64, 273-284.

Wicks, T.E., Smith, G.M., and Curran, P.J. (2002). Polygon-based aggregation of remotely sensed data for regional ecological analyses. *International Journal of Applied Earth Observation*, 4, 161-173.

Wilkinson, G.G. (1996). A review of current issues in the integration of GIS and remote sensing data. *International Journal of Geographic Information Systems*, 10, 85-101.

Woodcock, C. E., and Strahler. (1987) The factor of scale in remote sensing. *Remote Sensing of Environment*, 21, 311–332.

Zhang, Q., Wang, J. Gong, P. and Shi, P. (2003). Study of urban spatial patterns from SPOT panchromatic imagery using textural analysis.