Development of Forest Degradation Indicators from Long-term Trajectories of Multispectral Satellite Images, and their Projections into the Future under Climate Change, in Ontario, Canada

A Dissertation Submitted to the Committee on Graduate Studies

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the Faculty of Arts and Science

TRENT UNIVERSITY

Peterborough, Ontario, Canada

Environmental and Life Sciences Ph.D. Graduate Program

January 2024

ABSTRACT

Development of Forest Degradation Indicators from Long-term Trajectories of Multispectral Satellite Images, and their Projections into the Future under Climate Change, in Ontario, Canada

Md. Mozammel Hoque

Ontario forests are affected by natural and anthropogenic disturbances leading to forest degradation, which significantly impact local ecosystems, health, safety, and economy. This thesis develops a methodology for the continuous assessment, mapping, and monitoring of present and historic (1972–2020) forest disturbances, and future forest degradation trends and projections, using remote sensing data, ground measurements, and predictive models in an Ontario forested area.

After testing four supervised classification algorithms, support vector machine was found to be the most robust, consistent, and effective for land cover classification. Seven vegetation indices derived from Landsat and MODIS platforms were used to derive forest degradation indicators (FDIs), which were combined into one composite forest degradation indicator (CFDI) for each year, using the principal component analysis image fusion approach. The CFDI was the most informative indicator. The computed FDIs from available large multispectral image stacks were statistically related to historical climate variables. These relationships were used to project future FDIs related to climate variables derived from General Circulation Models through multiple linear regression models. Spatially-explicit maps of relevant climatic variables and of long-term historical forest degradation were developed from the LandTrendr trajectory analysis. Climate variables P, MA_1, MA_2, and CFDI were strongly correlated, allowing for the development of a model with a high coefficient of determination, R^2 (0.93), and low RMSE (0.28) to predict future values.

Forest disturbances (as CFDI) were also monitored from 1972–2020. Overall, these relationships allowed for to the creation of spatially-explicit, long-term historical forest degradation maps derived from the Landtrendr trajectory analysis.

Historical and future forest degradation maps identified the areas with projected high vulnerability to climate change, as well as the actual and potential changes in forest cover under climate change. The results indicated 2050 will experience an average temperature increase of 3.0°C, projected yearly decrease in precipitation of 109.5 mm, evapotranspiration increase of 73.0 mm, and moisture deficits of 28.47 mm (MA_1) and 37.60 mm (MA_2), leading to increased forest degradation.

Dedicated to

my parents

and all the members

of my family (Gusthi)

both present and future

ACKNOWLEDGEMENTS

In the name of God, The Most Gracious, The Most Merciful and with His blessing, I was able to complete this thesis.

Earning a PhD degree is a long and difficult journey, one filled with both excitement and pain. No one can walk this road alone and overcome its obstacles without sincere assistance from others. I know that acknowledging them here is not enough; however, it is a start.

First, I would like to express my deepest gratitude to my supervisor and mentor, Prof. Raul Ponce-Hernandez, for providing me with guidance and encouragement during the research and development of this thesis. In addition, I would like to thank my supervisory committee members, Dr. Carey Gibson and Dr. Trevor Jones, for their important input and suggestions that helped to improve this work. I would also like to thank my friends Katsuto Shimizu, Oumer Ahmed, Marsha Serville, Yeukai Katanda, Sohidul Islam, Anisur Rahman, and many others for continuously encouraging me to complete my thesis.

It is also my privilege to express my gratitude to Trent University, especially the office and technical staff in the Environmental and Life Sciences Department, for providing the necessary assistance and facilities for the completion of this thesis. Many thanks to NASA, USGS, Environment and Natural Resources Canada, Environment and Climate Change Canada, and the Canadian Centre for Climate Services for providing free open-source data. I am also very grateful to the Petawawa Research Forest for helping with personnel and providing the instruments to collect the data from the large and complex study area, which posed a number of challenges. In addition, I am thankful to the Ministry of Natural Resources and Forestry, Peterborough, Ontario for providing me with several land cover maps.

Finally, I wish to thank my wife, sons, brothers, sisters, and my parents for their love, support, and patience throughout my life. In particular, I want to express my gratitude to my youngest sister, Asst. Prof Monowara Begum, and her husband, Asst. Prof Nurur Rahman, for encouranging me during my studies. Without all their inspiration, this entire journey would have been meaningless.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND ACRONYMS

Artificial neural networks (ANN)

Canadian Centre for Climate Services (CCCS)

Climate moisture index (CMI)

Composite forest degradation indicator (CFDI)

Couple Model Intercomparison Project 5 (CMIP5)

Dynamical downscaling (DD)

Earth Observation (EO)

Empirical Bayesian kriging (EBK)

Enhanced vegetation index (EVI)

Evapotranspiration (ETo)

Forest degradation indicators (FDIs)

General circulation models (GCMs)

Global mean sea level (GMSL)

Greenhouse gas (GHG)

Intergovernmental Panel on Climate Change (IPCC)

Intrinsic random function kriging (IRFK)

Inverse distance weighted (IDW)

Kappa coefficient (K coefficient)

K-nearest neighbors (kNN)

Land cover/land use (LCLU)

Leaf area index (LAI)

Length of growing period (LGP)

Linear mixed model (LMM)

Mahalanobis distance (MahD)

Maximum likelihood classifier (MLC)

Net primary productivity (NPP)

Normalized burn ratio (NBR)

Normalized difference vegetation index (NDVI)

Moisture availability (MA)

Operational Land Imager (OLI)

Ordinary kriging (OK)

Overall accuracy (OA)

Precipitation (P)

Principal Component Analysis (PCA)

Producer's accuracy (PA)

Random forest (RF)

Regions of interest (ROIs)

South Saskatchewan River Basin (SSRB)

Support vector machine (SVM)

Tasseled cap brightness (TCB)

Tasseled cap greenness (TCG)

Tasseled cap transformation (TCT)

Tasseled cap wetness (TCW)

User's accuracy (UA)

Vegetation condition index (VCI)

Vegetation indices (VIs)

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Introduction

1.1.1 Background

1.1.1.1 Forest disturbance and degradation

Forest ecosystems, in particular, are tightly bound with climate both directly through the effects of precipitation, temperature, and related phenomena of evaporation and evapotranspiration, as well as directly controlling the moisture balance, and also indirectly through the effects of disturbances. Forest disturbances result in changes to the composition, structure, and character of a forest ecosystem (Jonson and Miyanishi, 2007), affecting its future health (Duinker, 2008). Natural disturbances, such as weather, geological forces, or biological changes, are one reason for an unbalanced ecosystem. Natural disturbances, which should increase in magnitude with a changing climate, can be caused by fire, insects, diseases, severe wind and ice storms, floods, volcanic activity, earthquakes, droughts, long-term freezing, hurricanes, landslides, etc., and can significantly damage an ecosystem, even killing flora and fauna.

Ecosystems are also affected by human activities, which are the cause of most forest disturbances, including clear cutting and the conversion of forest to other land uses, such as commercial and residential development, highways, and agriculture. Other human disturbances include uncontrolled wildfires, primarily caused by escaped or deliberately set fires, the introduction of non-native species, the elimination of species, and climate change. Slash-burning ("prescribed fire"), a well-established practice in Ontario (Lawson, 1990), exemplifies some of the harmful effects of anthropogenic activities. As well as emitting smoke and greenhouse gases (GHGs), it may have other negative repercussions, such a possible reduction and/or destruction of the long-term nutrient supply, effects on ecosystem structure and function, changes in biodiversity, and alterations in micro-organisms and organic matter in the soil (Fuglem, 1990).

Cohen et al. (2016) analyzed Landsat time series from 1985 to 2012 to show increasing nationwide forest disturbances in the US at a rate between 1.5% and 4.5% per year. Their results revealed a shift in disturbances; prior to the late 1990s, disturbances were primarily caused by anthropogenic forces, however, now most are caused by natural forces (e.g., climate, insects, and disease). Similarly, Ontario's forests are significantly influenced by these natural and human-made disturbances. In the Ontario boreal forest, in particular, insect outbreaks, forest/wildfires, permafrost melting, and logging could potentially worsen global warming. Additionally, industrial development could weaken both the forest's resistance and resilience to the intensifying negative consequences of global warming.

Climate change, including not only the rise in average air temperatures and in extreme weather events, such as the increased frequency, duration, and intensity of drought as well as summer heat waves, hurricanes, severe thunderstorms, windstorms, and ice storms (Duinker, 2008), will have a significant effect on the function, composition, and structure of forest ecosystems (Soja et al., 2007; Wulder et al., 2007). The rising intensity and prevalence of insect disturbances will also greatly impact forests (Wulder et al., 2007; Duinker, 2008). On a provincial scale, in Ontario, the supply of traditional forest products is expected to decline, accompanied by a negative impact on the forest industry and forest-dependent communities (Browne and Hunt, 2007). Climate change also threatens existing areas that protect specific natural features, species, and ecological communities (Scott and Lemieux, 2005). Besides the effect on biodiversity conservation, such changes can have dire consequences on nature-based recreation and tourism activities (Browne and Hunt, 2007).

Forest degradation leads to a "temporary or permanent decline in the net primary productivity (NPP) of forest ecosystems (measured in terms of $gC/m^2/yr$) and a deterioration in the density or structure of vegetation cover or its species composition" (Grainger, 1993), which reduces the forest's ability to adapt to a changing climate and provide ecosystem and livelihood services, particularly that of carbon sink. It

is a significant worldwide concern for several United Nations (UN) organizations and conventions (e.g., the UN Forum on Forests). However, the lack of an agreed common definition (and monitoring framework) (Lund, 2009) has impeded international recovery efforts (FAO, 2010).

1.1.1.2 Climate change

The World Meteorological Organization (WMO) defines climate as the average weather over an extended time period (for instance, three decades). A region's climate is the result of a complicated and interactive climate system with five main components: the atmosphere, hydrosphere, cryosphere, land surface, and biosphere. There are many biological, physical, and chemical processes involved in these interconnected elements. In the span of the past 4.5 billion years, the Earth's climate has gradually changed.

The Intergovernmental Panel on Climate Change (IPCC) defined climate change as "a change in the state of the climate that can be identified (e.g., by using statistical tests) by changes in the mean and/or the variability of its properties, and that persists for an extended period, typically decades or longer. Climate change may be due to natural processes or external forcings, or to persistent anthropogenic changes in the composition of the atmosphere or land use." Due to the naturally occurring variability in the planet's climate over the short-term, irrefutably identifying the long-term reasons for contemporary climate change poses difficulties; however, the global anthropogenic impact, which is expected to continue, is becoming increasingly pronounced (IPCC, 1996), leading to perceptible changes in air temperatures, precipitation, and extreme weather (IPCC, 2007).

The accumulation of GHGs (e.g., carbon dioxide, methane, etc.) in the atmosphere (IPCC, 2007) causes global warming. This allows shorter wavelength radiation (incoming solar) to reach the surface of the planet while absorbing longer wavelengths (outgoing terrestrial), which is, in part, radiated back to the Earth. GHGs in the atmosphere come primarily from burning fossil fuels (80%), while changes in land use and cover, especially forest fires, forest degradation, and deforestation, contribute the rest (20%)

(Stern, 2007). Anthropogenic activities have increased atmospheric GHG concentration, and as a result, the mean annual global temperature is predicted to rise from 1.8–4°C during the 21st century, leading to observable and harmful changes, such as variations in rainfall patterns, a rise in sea level due to ice and snow melt, and an increase in the intensity and frequency of extreme weather events (IPCC, 2007). Therefore, time is short to stabilize and reduce GHG emissions to avoid their "devastating impacts on our planet" (Environment, Heritage, and Local Government, 2007). In this light, it is crucial to develop reliable methods for monitoring and predicting present and future changes as well as mitigating their effects.

A synopsis of climate change projections for Ontario, Canada, indicated a 4.7 to 7.9^0C rise in average annual air temperature will occur in the basins of Lake Huron and the Ottawa River, when compared to the record of the years from 1971–2000 (McDermid et al., 2015). Climate warming is expected to lead to longer growing seasons (1–30% greater in length by 2100), as well as increased forest fire frequency and pest outbreaks in Canada. The appearance and functioning of ecosystems, by the end of this century, may also be very different when compared to today. Since the study area, which includes Algonquin Park (a designated protected area) is very important in conservation and public interest, it is crucial to have the information provided by updated climate warming projections at a scale useful for the management of the park. These will help overcome the previous disparity between climate model projections at very coarse scales and potential effects at fine scales in both terrestrial and aquatic ecosystems (Maclean et al., 2017; Nadeau et al., 2017).

1.1.1.3 Forests and climate change

Forests are crucial in maintaining the delicate relationship between natural ecosystems. As the climate and forests are interrelated, they affect each other in both direct and other ways. Unfortunately, forests are expected to face significant pressure in the next several decades resulting from the effects of rising GHG

concentrations and climate change (IPCC, 2013), potentially disrupting the important ecological, economic, social, recreational, and aesthetic services they provide (Eastaugh, 2008). From the perspective of climate change, all human activities that act to decrease the present carbon stock (carbon carrying capacity) in natural forests need to be included as the primary drivers and pressures on forest degradation (Mackey et al., 2008).

Scientists generally agree that a higher concentration of CO_2 in the Earth's atmosphere is an important driver of global warming (IPCC, 2013), as well as that forests are critical for mitigating climate change (Nabuurs et al., 2007). Forests contain stores of approximately 45% of terrestrial carbon and, according to estimations, half the terrestrial primary productivity can be attributed to forests (Bonan, 2008a). As a result of this productivity, large absorptions of atmospheric CO_2 occur. In the northern hemispheres of the planet, forests are great contributors to a significant portion of this carbon sink, in particular the mid- and high-latitude ecosystems (Goodale et al., 2002; Kurz et al., 2008b). It is projected that these high latitude forests will be greatly affected by climate change, which can both positively and negatively influence the carbon balance (Kurz et al., 2008b). For instance, a warmer climate will likely positively affect forest productivity, such as with longer growing seasons. However, it is also projected that a warmer climate will increase the impact of forest disturbances of Earth's ecosystems (Seidl et al., 2014). In forests in Europe, fires, winds, and insect outbreaks have already produced signs of saturation of the carbon sink (Nabuurs et al. 2007). Additionally, it is likely that global warming will have further effects on the temporal and spatial dynamics of energy flows, as well as pest outbreaks, both in terms of the intensity and range (Battisti, 2008; Hicke et al., 2012). However, insect outbreaks are usually not included in large-scale carbon modelling, the effect of which is the overestimation of forests' carbon sequestration capabilities (Kurz et al., 2008b; Hicke et al., 2012). There are also significant uncertainties related to the quantitative effects of insect outbreaks on the carbon balance (Clark et al., 2010; Hicke et al., 2012). It is

crucial, therefore, to develop ways of engaging in the monitoring of insect disturbances and quantify their impact on carbon balances.

The largest portion of Canadian boreal forests is located in Ontario (Ontario Ministry of Natural Resources, 2008b), where, since the mid-20th century, forests have been negatively impacted by human activities, which are projected to change the composition and function of ecosystems to a large extent (Schindler, 1998). Currently, a shift is being registered in the boreal forest from a net carbon sink to a net carbon source, primarily as a result of the increased intensity, area, and frequency of fires resulting from increased temperatures (Flannigan et al., 2005a, 2005b; Soja et al., 2007). The expected impact of climate change in boreal forests could be increases in ignitions from lightning, fire season length, and fire weather severity (Flannigan et al., 1998; Flannigan et al., 2001; Stocks et al., 1998, as cited in Amiro et al., 2003) and greatly threaten life and property. Climate change, including the rise in extreme weather events, such as the increased frequency, duration, and intensity of drought as well as summer heat waves, hurricanes, severe thunderstorms, windstorms, and ice storms (Duinker, 2008), will have a significant effect on the function, composition, and structure of forest ecosystems (Soja et al., 2007; Wulder et al., 2007). The rising intensity and prevalence of insect disturbances will also greatly impact forests and their health (Wulder et al., 2007; Duinker, 2008). The supply of traditional forest products is expected to decline, accompanied by a negative impact on the forest industry and forest-dependent communities (Browne and Hunt, 2007). Climate change also threatens existing areas that protect specific natural features, species, and ecological communities (Scott and Lemieux, 2005). Besides the effect on biodiversity conservation, such changes can have dire consequences on nature-based recreational activities and tourism (Browne and Hunt, 2007). As this social-ecological system is crucial and susceptible to climate change, it is important to identify effective methods for assessing and mapping the spatial distribution of forest disturbance and degradation, as well as identifying strategies for mitigation and adaptation.

1.1.1.4 Forest disturbance and degradation mapping

Landscape disturbance mapping and monitoring is greatly dependent on scale. Therefore, the landscape processes and patterns discernable with a remotely sensed image source rely on the target (e.g., single tree or stand replacing disturbance) and the spatial, spectral, temporal, and radiometric characteristics of the image's source (Perera and Euler, 2000). When detecting landscape change and disturbance, the primary objective is data comparison from a series of points in time by 1) controlling for all external factors due to differences in variables not of interest and 2) evaluating the actual changes resulting from the variable of interest (Lu et al., 2004). Owing to the vast number of time series approaches and detection algorithms created for change and disturbance detection in imagery, selecting and implementing the most appropriate method is crucial (Coppin et al., 2004; Lu et al., 2004). The 11 categories of change detection techniques developed by Singh (1989) can be grouped into five approaches (Coops et al.; in Wulder and Franklin, 2007):

(i) Image algebra (differencing, subtracting, or ratioing) of two or more images;

(ii) Regression or correlation, where the developed model predicts or compares spectral responses in a series of images;

(iii) Statistical techniques that compute statistical components that are then compared for temporal changes;

(iv) Classification comparisons in which images are classified separately and these classifications are then compared; Recent advances in this field incorporate pixel-based classification, object-based classification, and machine-learning algorithms (e.g., support vector machines);

(v) Tools that analyze images and other datasets within a geographic information system (GIS) (Examining the statistical relationships between the spatial patterns of a select group of indicators and climate layers in raster format could be useful in obtaining information on change over time).

1.1.2 Statement of the research problem

While a significant area of Ontario forests is subject to disturbance or degradation, there are few studies on forest degradation and climate change in Ontario. This issue is compounded by the fact that no comprehensive definition of forest degradation in terms of detectable and measureable indicators exists. Additionally, there are no highly accurate existing land cover/land use classification maps for the area. The investigation of the compound effects of climate change requires a comprehensive set of tools and procedures for the long-term assessment, mapping, and monitoring of forest degradation. Remote sensing has provided an extremely suitable range of spatially explicit tools for such investigations. Thus, the design of a methodological framework, based on a carefully selected set of forest disturbance and degradation indicators derived from existing and available data, imagery, and modelling tools is necessary to address this assessment problem. The development of procedures for spatially-explicit forest degradation indicator estimation, together with spatially-explicit projections of climate change, will help elucidate the causality between climate change variables and the incidence of disturbances, discounting for anthropogenic factors, through spatial analysis with demographic layers. The ensemble of key tools and procedures for the mapping and continuous monitoring of the spatial pattern of forest disturbances and degradation over time would allow for evidence-based management and preparedness as well as suitable and timely policy interventions.

A composite of the main technical issues that are part of the research problem in this thesis is shown graphically in Figure 1.1 below.

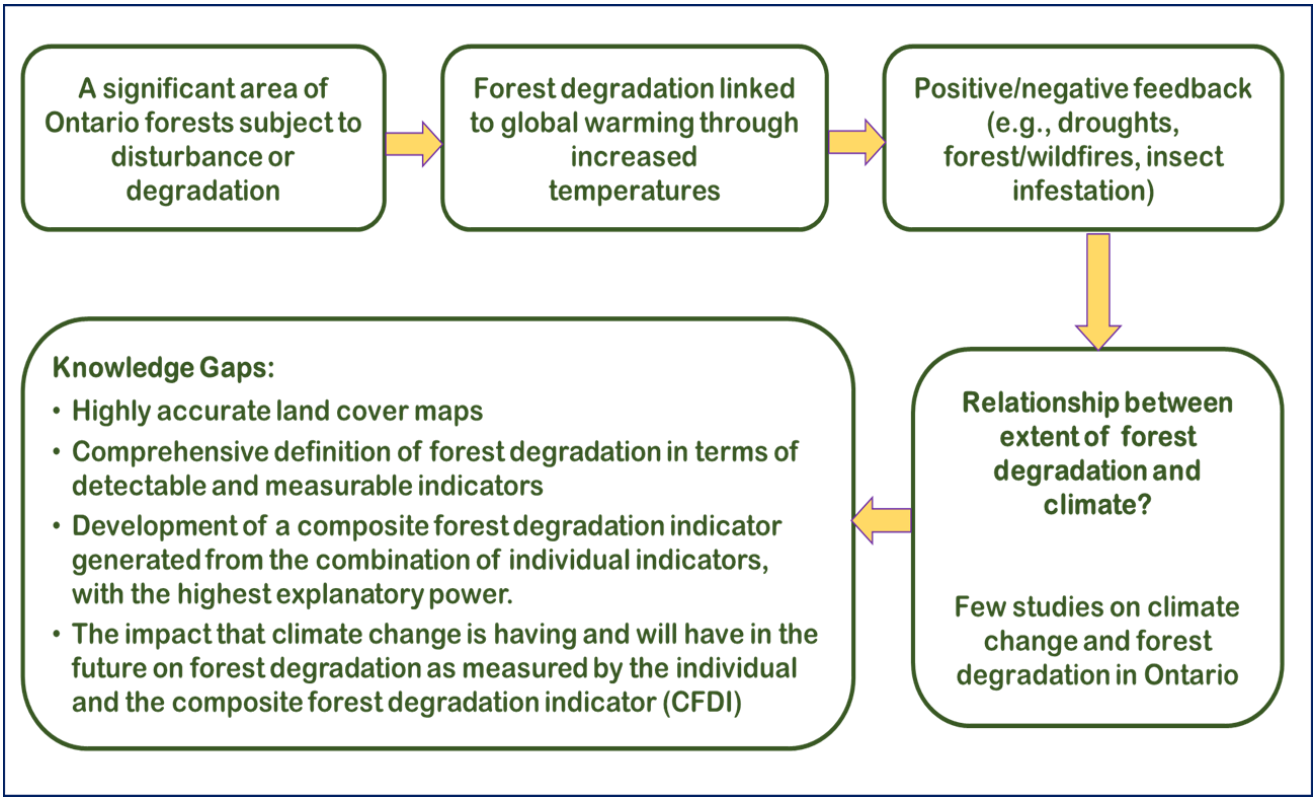


Figure 1. 1: Graphical representation of the research problems

1.1.3 Specific research questions

- What is the most common definition of forest degradation used by international bodies and national forest experts that can be applicable to Ontario forests?
- Can degraded forests be distinguished and accurately mapped from non-degraded forests using a range of existing satellite imagery products and derived computed indicators of forest degradation?
- What is the spatial pattern of forest cover change in the selected area of Ontario?
- What driving forces affected forest cover change and degradation from 1972–2020?
- How will present forest degradation processes be affected by climate change?
- What is the spatial pattern of degraded forests in Algonquin Park expected for 2050 due to climate change?

1.1.4 Objectives of the study

1.1.4.1 Overall objective

The overall objective of this research is the development of a methodological framework for the continuous assessment, mapping, and monitoring of present and historic forest disturbance and degradation trends and their projections under climate change, using remote sensing technology, existing ground measurements, and predictive models, such as regional circulation models (RCMs). Specifically, the research will identify, test, and establish procedures for the continuous monitoring of forest disturbance and degradation throughout a selected portion of Ontario, and establish, to the extent possible, the nature and pattern of climate change impacts on these Ontario forests.

1.1.4.2 Specific objectives

- To examine and evaluate the accuracy of conventional and relatively new machine learning algorithms for highly accurate land cover classification based on Landsat imagery of central and northeastern Ontario, Canada.
- To identify a suitable set of forest degradation indicators (FDIs) derived from multispectral satellite imagery that can be used in the region and then compared with those defined by international standards.
- Develop an innovative composite forest degradation indicator (CFDI) to examine the relationship with climate variables and future (2050) projection of forest degradation under climate change.
- Long-term trajectory analysis using the FDIs, including CFDI, and generating disturbance and degradation maps for 1972–2020.
- To use the established methodological framework to assess the trends in forest cover change in the region during the longest period possible (e.g., 1972–2020), and the relationship between the trends of forest cover change and their underlying causative factors, particularly their relationship to climate change.

- To find out the dependence of disturbance or degradation on or at least its degree of correlation with climate variables.
- To predict the expected change in forest condition (degradation) with climate change in the forest area for the region in the future (e.g., year 2050).
- To determine the consequences of projected forest degradation with climate change in the region of study relative to GHG concentrations and fluxes. This will hopefully serve as the basis for the development of appropriate climate change mitigation and adaptation strategies for forest ecosystem and biodiversity management for the region in the future.

1.1.5 Approach and methods

This research used geospatial, remote sensing, computer modelling, statistical, and indicator approaches, as well as existent forest resource inventory data, to achieve its objectives. These tools were appropriately combined into a methodological framework after identification of the best options in terms of modelling, remote sensing and statistical tools, and procedures. As remote sensing can provide a great deal of pertinent environmental data over a broad range of spatial and temporal scales, it is likely to play a crucial role in providing environmental condition indicators. A set of indicators of forest disturbance and degradation were investigated, identified, and defined after an intense literature review. Direct methods of remote sensing were used to assess canopy damage using time series data to detect and map degraded forests, while indirect methods were employed to estimate the forest area affected by forest degradation detecting human infrastructure and land use change as “proxies.” A combination of both direct and indirect methods were also used to assess forest degradation.

Parametric and nonparametric statistical techniques were employed for land cover classification and to examine the possible relationships between meteorological variables that indicated climate change and the spatial and temporal distributions of indicators of land degradation. Predictive linear multiple

regression models were derived based on statistically significant relationships. Analysis of variance and multivariate tests were incorporated into the data analysis, as required, and applied to the indicators of forest degradation deemed useful in this context. Machine learning algorithms (e.g., support vector machine) were compared to pixel-based classification to determine the land cover classes involved in the results of the trajectory analysis by any of the selected and commonly accepted algorithms in the scientific literature (e.g., LandTrendr, BFAST, etc.). The usefulness of wavelet transformations and principal component analysis was investigated for the fusion of multiple sources of data and platforms (e.g., multispectral with Landsat and MODIS data), when available.

1.2 Study area, data, and preprocessing

1.2.1 Study area

The area selected for this study, in terms of spatial extent, spans from 45°8’15.32’’N to 46°16’12.38’’N and from 76°37’48.75’’W to 78°20’16.73’’W, defining a total of 10,000 km^2. This study area (Figure 1.2) contains a wide range of forest and non-forest land cover types, as it is typical of forests in the boreal transitional zone, including Algonquin Park. For a more detailed description of the study area, please see Chapter 2 (Section 2.2.1).

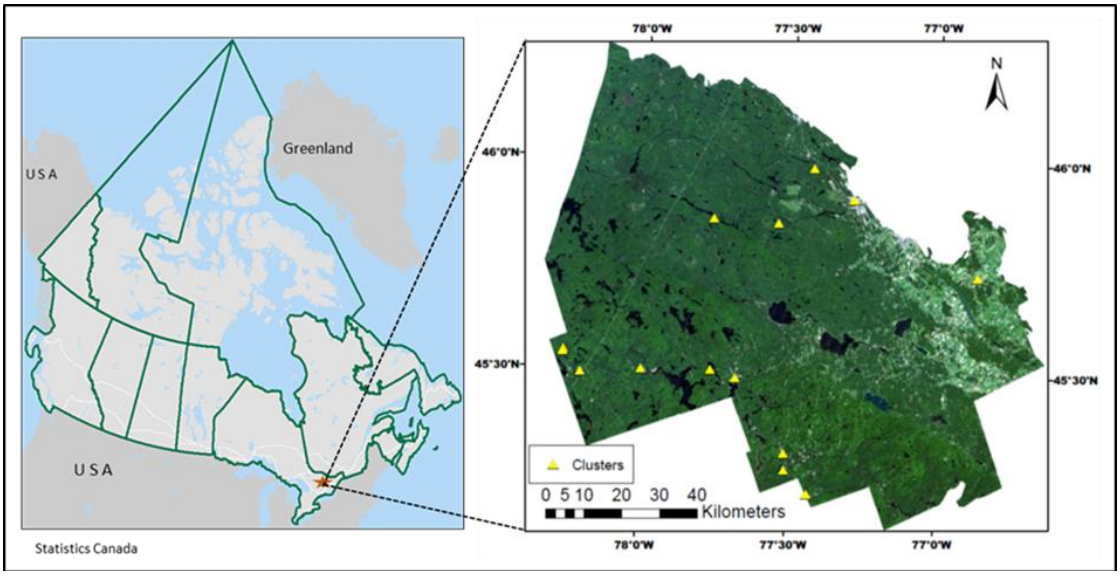


Figure 1. 2: Study area with clusters in central and northeastern Ontario, Canada

1.2.2 Data

1.2.2.1 Landsat imagery

Landsat time series stack (LTSS)

The study area lies entirely within Landsat path 17 (MSS 18), row 28. One of the research objectives was to use the entire Landsat image record (1972–2020) to monitor historic, present, and future forest disturbance and degradation trends. This work was divided into several parts, guided by the time ranges within which satellite imagery were available and of good quality: Landsat 1–3 Multi-Spectral Scanner (MSS) (1972–1982), Landsat 4 & 5 MSS (1983–1987/92/93), Landsat 4 &5 Thematic Mapper (TM) (1982–2012), Landsat 7, Enhanced Thematic Mapper Plus (ETM+) (1999–present), and Landsat 8 Operational Land Imager/Thermal Infrared Sensor (OLI/TIRS) (2013–present).

The Landsat satellite images for 1972–2020 were acquired from USGS websites: glovis.usgs.gov and earthexplorer.usgs.gov. Landsat 1–5 MSS imagery from 1972–1983 were downloaded, and the collection of the same images was repeated when they were released in tier 2 in 2018, which had classification accuracy far better than older images. Landsat TM and ETM+ data were acquired for the period from 1984–2013, as well as collected Surface Reflected data (collection-1, level-2), using the USGS "On Demand" option. Landsat 8 data for 2013–2020 were also acquired, and Surface Reflected data (collection-1, level-2) were collected using the same technique.

A summer season window of June 21 to the end of September was used for the satellite image selection period. From this time window, 3–5 images per year were downloaded, and the best one was selected after layer stack and pre-processing. For each year, one cloud-free or mildly cloudy contaminated image was chosen for the building of the LTSS; if there was no image that met these criteria, several partly cloudy (<10% CC) images with a similar phenology were employed as a replacement (see Table 1.1).

Table 1. 1: Detailed information for the Landsat satellite images (1972–2020)

Acquisition Date			Julian	Landsat	Landsat	Cloud
Year	Month	Day	DOY	Platform	Sensors	Cover
2020	08	12	225	LC8	OLI_TIRS	0.2%
2019	06	23	174	LC8	OLI_TIRS	0.1%
2018	09	08	251	LC8	OLI_TIRS	0.5%
2017	08	28	240	LE7	ETM+	10%
2016	06	30	182	LC8	OLI_TIRS	0%
2015	09	16	228	LC8	OLI_TIRS	1%
2014	06	09	160	LC8	OLI_TIRS	0%
2013	09	26	269	LC8	OLI_TIRS	1%
2012	10	01	275	LE7	ETM+	6%
2011	10	07	280	LT5	TM	0%
2010	05	21	141	LE7	ETM+	0%
2009	06	27	178	LT5	TM	2%
2007	08	01	213	LE7	ETM+	4%
2006	08	30	242	LE7	ETM+	0%
2005	09	04	247	LT5	TM	1%
2004	08	24	237	LE7	ETM+	0%
2003	05	18	138	LE7	ETM+	0%
2002	06	24	175	LT5	TM	0%
2001	08	24	236	LT5	TM	0%
2000	09	22	266	LT5	TM	4%
1998	08	16	228	LT5	TM	12%
1996	08	10	223	LT5	TM	4%
1995	06	21	172	LT5	TM	0%
1994	07	04	185	LT5	TM	0%
1993	07	01	182	LT5	TM	0%
1992	08	15	228	LT5	TM	2%
1991	08	29	241	LT5	TM	0%
1990	07	25	203	LT5	TM	2%
1988	07	03	185	LT5	TM	1%
1987	06	15	166	LT5	TM	5%
1986	08	31	243	LT5	TM	0%
1985	08	12	224	LT5	TM	2%
1984	06	22	174	LT5	TM	1%
1981	08	18	230	LM2	MSS	0%
1980	07	18	200	LM2	MSS	0%
1979	06	18	169	LM2	MSS	0%
1978	07	02	182	LM3	MSS	0%
1976	10	01	275	LM2	MSS	0%
1975	09	28	271	LM1	MSS	0%
1973	10	08	281	LM1	MSS	0%
1972	08	20	233	LM1	MSS	0%

Note: Some years are missing since there is no available and/or usable Landsat image data for those dates.

1.2.2.2 MODIS imagery/MODIS NPP data products

MODIS gap-filled yearly NPP products (MOD17A3HGF-V061) with a spatial pixel resolution of 500m were acquired from the NASA archive: modis.gsfc.nasa.gov. NPP products were obtained based on the difference of the gross primary production (GPP) and the maintenance respiration (MR). The GPP algorithm applies the concept of light use efficiency from both MODIS LAI/FPAR product and also NASA GMAO meteorological data (Running and Zhao, 2015). NPP data from 2001–2020 were acquired and applied in this study to create the composite forest degradation indicator (CFDI) and employ the well tested trajectory analysis algorithm known as "LandTrendr" to determine the long-term forest disturbance and degradation in the study area.

1.2.2.3 Ground measurements, ground truth, and field sampling

All ground data were collected from July to September 2018 (the area's growing seasons). Accessibility was an important and determining factor for data collection in terms of the number of samples and sample spatial distribution. A randomized cluster sampling design was adopted for this heterogeneous study area, as it was difficult to collect field data due to the limitations of access roads and the risk posed by wildlife. For more information about field sampling, see Chapter 2 (Section 2.2.2).

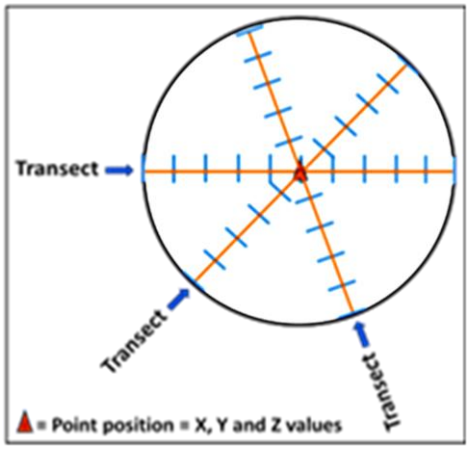


Figure 1. 3: Collection of a point data from the field using transect

1.2.2.4 Historical and projected climate data

1.2.2.4.1 Historic climate data

Historical climate (daily and monthly) data for the period from 1972–2020 were acquired from the official climate data sources of Environment and Natural Resources Canada. In addition, historical Canadian climate normals data were garnered from the Environment and Climate Change Canada (ECCC) official

climate data portal. Two sets of 30-year climate normals data (1971–2000 and 1981–2010) were collected from 51 weather stations inside and surrounding the study area.

1.2.2.4.2 Projected climate data: General circulation model (GCM) and downscaled RCM data

Monthly observations of maximum temperature (T_{max}), minimum temperature (T_{min}), mean temperature (T_{mean}), and monthly total precipitation (P) for 2050 were acquired from the Canadian Centre for Climate Services (CCCS) data portal (https://climatedata.ca). These data were already calculated by the CMIP5 climate modelling experiment dataset and statistically downscaled and bias adjusted using the "BCCAQv2" method under representative concentration pathway (RCP) 8.5. An ensemble of 24 GCMs was used. RCP8.5 was selected as it is considered the maximum concentration of greenhouse gases (GHGs) causing a forcing of 8.5 watts m^{-2}. This would occur if there is minimal to no mitigation of GHGs ("worst case scenario").

1.2.3 Preprocessing

An image stack was created for all downloaded images by selecting all the required bands using the Layer Stacking tool from ENVI^{TM} 5.3 software. The images from 1972–2020 were stacked for each year for preprocessing.

Landsat image preprocessing included the detection and restoration of bad lines, radiometric correction, gap filling, cloud and cloud shadow removal, as well as composite filtering. To correct for aerosol and cloud effects, radiometric calibration and atmospheric correction were conducted with the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm (Masek et al., 2006). Clouds and their shadows were automatically screened using the function of mask (FMASK) algorithm (Zhu and Woodcock, 2012).

Images with thin clouds and haze creating blurriness were cleaned with the Dark Object Subtraction tool in ENVI. All the images of the study area were also cleaned using Adaptive (Additive) filter "Local

Sigma" and the Convolution (Multiplicative) filter "Median" to remove the "salt-and-pepper" noise (speckle, spike, sharpness, etc.).

Landsat 1–5 MSS images were corrected using Exelis Visual Information Solutions (ENVI ^{TM} 5.3). As it is not possible to calibrate or atmospherically correct old MSS data in a reliable manner, every scene from the 1975-era MSS (Landsat 1–5 MSS) archive was fixed radiometrically in agreement with the ETM+ image (as in Hall et al., 1991).

Landsat 4–5 TM and Landsat 7 ETM+ images were atmospherically corrected to a surface reflectance image with the LEDAPS algorithm (Masek et al., 2006, 2013; Schmidt et al., 2013; USGS, 2017, 2020), which employs the Second Simulation of the Satellite Signal in the Solar Spectrum (6S) radiative transfer model as described by Vermote et al. (1997). LEDAPS transforms digital numbers into surface reflectance values and also conducts radiance calibration, atmospheric correction, and top-of-atmosphere reflectance conversion. The combination of atmospheric correction and normalization increases the accuracy of the spectral reflectance trajectory model in comparison with absolute correction (Schroeder et al., 2006). For remote sensing image processing, the reconstruction of gaps is a crucial concern. The SLC gap filling approach suggested by Romero-Sanchez et al. (2015) consisting of the Geostatistical Neighborhood Similar Pixel Interpolator (GNSPI) was used in the reconstruction of the imagery stack of Landsat 7 ETM+. In this study, the ETM+ SLC-off images (2004–2017) were fixed by employing the Landsat Gap fill IDL Model of ENVI (ENVI 5.3.) with the GNSPI algorithm (Zhu et al., 2012).

Landsat 8 OLI surface reflectance products were obtained already corrected by the LaSRC algorithm (USGS, 2016, 2020). These corrections increased the radiometric resolution and improved signal-to-noise ratio (Vermote et al., 2016; USGS, 2019). As a result, Landsat 8 surface reflectance product is far superior to the makeshift Landsat 5/7 LEDAPS product (Vermote et al., 2016; USGS, 2019). Fmask 4.0 (beta version), which automatically detects water, snow, clouds, and cloud shadows, was employed for Landsat

4–8 and Sentinel-2 images, which was developed based on Fmask and MFmask algorithms version 3.3 (Qiu et al., 2017, 2018, 2019). Therefore, for this study, Fmask 4.0 (beta version) was employed to first detect and then remove clouds and cloud shadows from Landsat 4–8 images, most of which were "cleaned" by this algorithm. For any images that were not entirely cleaned, the Landsat Pixel Quality Assessment band (Pixel_qa band) was employed to remove the cloud and cloud shadows.

1.3 Overall methodological workflow

The overall methodology for this project consisted of a series of analytical steps that, when put in sequence, made up a methodological workflow, which could be organized as a flowchart. The workflow for the methodology in this project is shown in Figures 1.4 and 1.5.

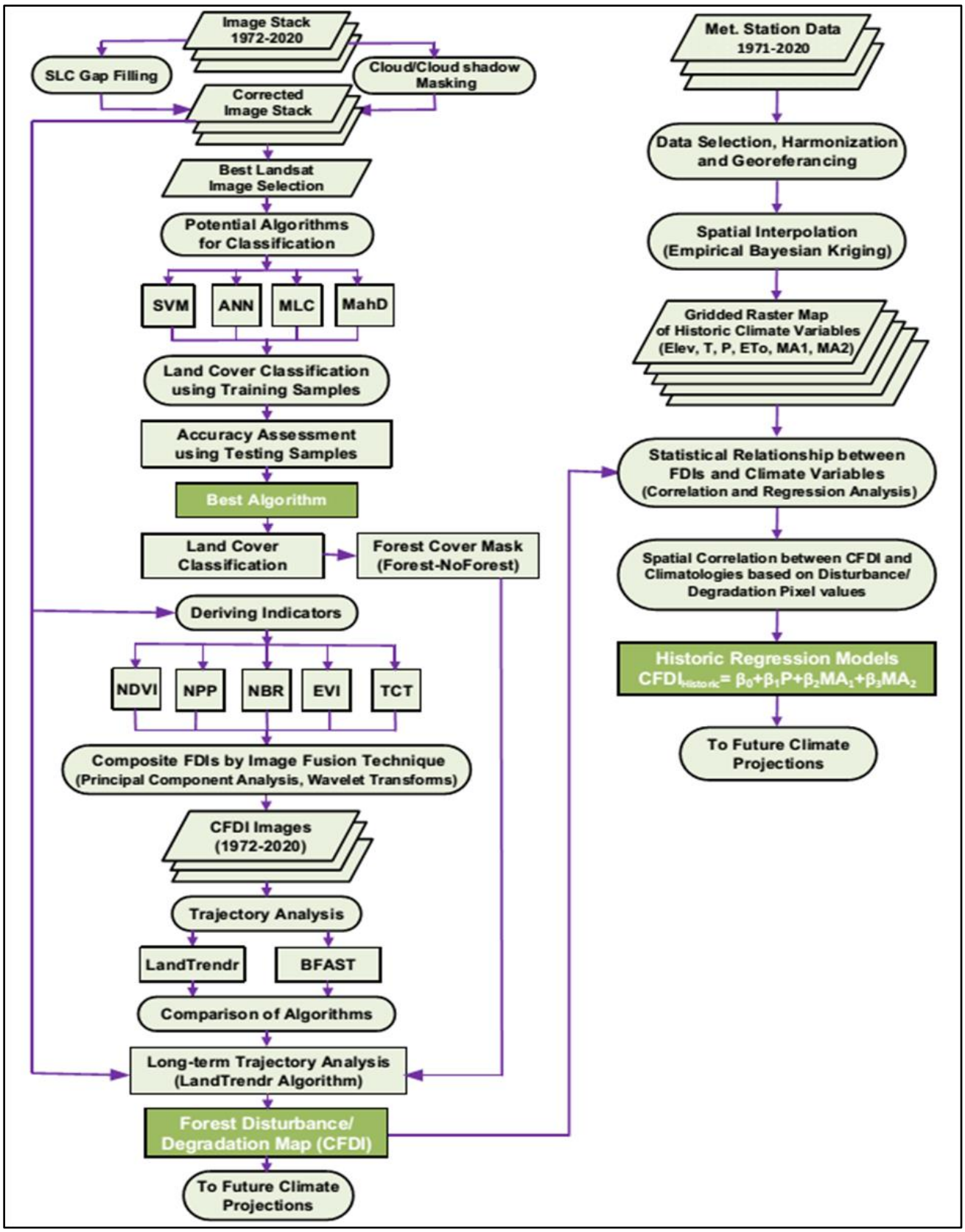


Figure 1. 4: Methodological workflow for forest disturbance and degradation assessment

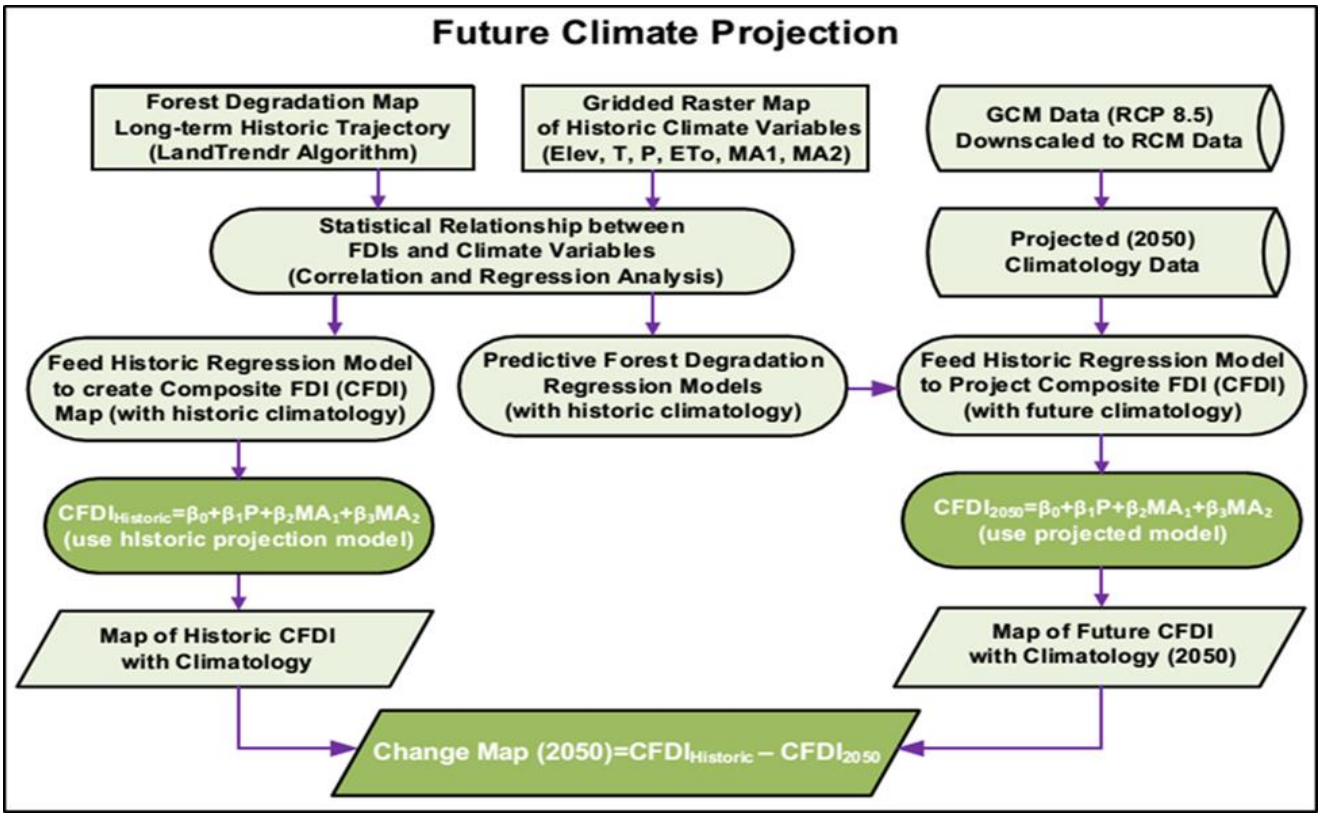


Figure 1. 5: Methodological workflow for projected future forest degradation under climate change

1.4 Dissertation organization

This thesis is presented in such a way that each chapter (with the exception of the general introduction and conclusions) may be read independently. The research chapters all match the publication (submitted) or prepared for publication versions, with changes only to formatting to maintain a consistent style through the thesis. The thesis is composed of six chapters, four of which are research chapters (one was submitted to the Canadian Journal of Remote Sensing (CJRS) after review, and two abstracts were accepted for a symposium and a conference, including one extended abstract of 2000 words.)

Chapter 1 presents the general introduction, background/literature review, brief description of forests and climate change**,** problem statement, objectives of the study, approach and methods, study area and data collection, overall methodological flow chart, and thesis/dissertation organization.

Chapter 2 provides information about the field sampling design, study area and data collection, remote sensing data, methodological workflow of the chapter, land cover, image classification, and imbalanced training dataset. This chapter examines the performance of four classifiers with imbalanced ground truth data and also assesses the sensitivity of Support Vector Machine (AI) classification performance to training class with imbalanced data. Under the title "Evaluation of the accuracy of multispectral land cover classification algorithms: Comparison of Machine Learning, Maximum Likelihood, and Mahalanobis Distance Algorithms," this chapter was submitted to CJRS for publication.

Chapter 3 introduces vegetation indices that are applied as FDIs in this study. It also describes the innovative CFDIs derived from Landsat and MODIS image data using principal component analysis, and the use of the LandTrendr algorithm for long-term trajectory analysis.

Chapter 4 details the relationship between FDIs, CFDIs, and climate variables. It provides a brief background of climate normals data, evapotranspiration, moisture availability, length of growing period, drought, and empirical Bayesian kriging (EBK) as a mechanism for optimal spatial interpolation of point-data, forest disturbance and degradation, FDIs, Ontario forests, drought, and LandTrendr. It also involves a description of the historical multiple linear regression model, development of predicted (2050) multiple linear regression models, and the creation of the historical CFDI gridded maps of climatology (1972–2020).

Chapter 5 shows the methodological workflow of future climate projections. It also details the climate change impacts, emission scenarios, the impacts of drought under a changing climate, Ontario forests and climate change, forest disturbance and degradation, projection of future climate scenarios, GCMs, RCMs, and RCPs. In addition, it provides information about the climate change impact on the study area including Algonquin Provincial Park in 2050 through the creation of projected and change CFDI gridded maps of climatology (2050) that identify the most vulnerable forest degradation areas under the climate change

scenario. Under the title of "Projection of Future (2050) Forest Degradation Indicators under a Changing Climate in Central and Northeastern Ontario, Canada," an abstract of this chapter was accepted for presentation at the 44th Canadian Symposium on Remote Sensing and by The International Environmetrics Society (2023).

Finally, chapter 6 presents the overall conclusions and major findings of this research.

References

Amiro, B.D., MacPherson, J.I., Desjardins, R.L., Chen, J.M., and Liu, J. 2003. Post-fire carbon dioxide fluxes in the western Canadian boreal forest: Evidence from towers, aircraft and remote sensing, Agriculture and Forest Meteorology, 115, pp.91–107.

Battisti, A. 2008. Forests and climate change–lessons from insects. IFOREST BIOGEOSCIENCES AND FORESTRY, 1, pp.1–5.

Bonan, G.B. 2008a. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. Science, 320, pp.1444–1449.

Browne, S.A, and Hunt, L.M. 2007. Climate change and nature-based tourism, outdoor recreation, and forestry in Ontario: potential effects and adaptation strategies, climate change research report. Ontario Ministry of Natural Resources, Toronto.

Canadian Centre for Climate Services: (https://climatedata.ca)

Clark, K.L., Skowronski, N., and Hom, J. 2010. Invasive insects impact forest carbon dynamics. Global Change Biology, 16, pp.88–101.

Cohen, W.B., Yang, Z., Stehman, S.V., Schroeder, T.A., Bell, D.M., Masek, J.G., Hung, C., and Meigs, G.W. 2016. Forest disturbance across the conterminous United States from 1985–2012: The emerging dominance of forest decline. Forest Ecology and Management, 360, pp.242–252.

Coops, N.C., Wulder, M.A., and White, J.C. 2007. Identifying and describing forest disturbance and spatial pattern: Data selection issues and methodological implications, pp.31–61 in Wulder and Franklin, (eds.) Understanding Forest Disturbance and Spatial Pattern: Remote sensing and GIS approaches. Taylor and Francis.

Coppin, P., Lambin, E., Jonckheere, I., Nackaerts, K., and Muys, B. 2004. Digital change detection methods in ecosystem monitoring: A review. International Journal of Remote Sensing, 25, pp.1565−1596.

Duinker, P.N. 2008. Global climate change. Sustainable forest management network.

Eastaugh, C. 2008. Adaptations of forests to climate change: a multidisciplinary review. IUFRO Secretariat, Vienna.

Environment and Climate Change Canada: https://climate.weather.gc.ca/climate_normals/index_e.html)

Environment and Natural Resources Canada. 2018. Canadian Climate Normals 1981–2020. Station Data. Government of Canada, Environment and Natural Resources, Ottawa, Ontario. http://climate.weather.gc.ca/climate_normals.

Environment and Natural Resources Canada: https://climate.weather.gc.ca/historical_data/search_historic_data_e.html)

Environment, Heritage and Local Government (Delegation of Ministerial Functions) Order 2007.

FAO. 2010. Global forest resources assessment 2010. FAO Forestry Paper 163. Food and Agriculture Organization, Rome, Italy.

Flannigan, M.D., Amiro, B.D. Logan, K.A. Stocks, B.J., and Wotton, B.M. 2005a. Forest fires and climate change in the 21st century. Mitig Adapt Strategies Glob Change, 11, pp.847–859.

Flannigan, M.D., Logan, K.A., Amiro, B.D., Skinner, W.R., and Stocks, B.J. 2005b. Future area burned in Canada. Climate Change, 72, pp.1–16.

Fuglem, P. 1990. Prescribed Burning in British Columbia. B.C. Forest Service.

Goodale, C.L., Apps, M.J., Birdsey, R.A., Field, C.B., Heath, L.S., Houghton, R.A., Jenkins, J.C., Kohlmaier, G.H., Kurz, W., Liu, S., Nabuurs, G.-J., Nilsson, S., and Shvidenko, A.Z. 2002. Forest carbon sinks in the Northern Hemisphere. Ecological Applications, 12, pp.891–899.

Grainger, A. 1993. Controlling Tropical Deforestation, p. 310. Earthscan Publications Ltd, London.

Gudmundsson, L., Bremnes, J.B., Haugen, J.E., and Engen-Skaugen, T. 2012. Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods. Hydrol. Earth Syst. Sci., 16, pp.3383–3390.

Hall, F.G., Strebel, D.E., Nickeson, J.E and Goetz, S.J. 1991. Radiometric rectification: Toward a common radiometric response among multidate, multisensor images. Remote Sens. Environ., 35, pp. 11–27.

Hicke, J.A., Allen, C.D., Desai, A.R., Dietze, M.C., Hall, R.J., Hogg, E.H., Kashian, D.M., Moore, D., Raffa, K.F., Sturrock, R.N., and Vogelmann, J. 2012. Effects of biotic disturbances on forest carbon cycling in the United States and Canada. Global Change Biology, 18, pp.7–34.

IPCC. 1996. Climate Change 1995. The Science of Climate Change. J.T. Houghton, L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell (Eds). Cambridge University Press, UK.

IPCC. 2007. Climate change 2007: synthesis report. Cambridge University Press, Cambridge.

IPCC. 2013. Intergovernmental Panel on Climate Change. Climate Change 2013: The Physical Science Basis.

Jonson, E.A., and Miyanishi, K. 2007. Plant disturbance ecology. Academic Press, London.

Kurz, W.A., Stinson, G., and Rampley, G. 2008b. Could increased boreal forest ecosystem productivity offset carbon losses from increased disturbances? Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363, pp.2259–2268.

Lawson, B. D. 1990. "Where There Is Fire, There's Smoke: A Global View of B.C.'s Prescribed Burning," Presentation at Panel Discussion on Smoke Management, Southern Interior Fire Management Committee, Cranbrook, B.C., Canada.

Lu, D., Mausel, P., Brondizio, E., and Moran, E. 2004. Change detection techniques. International Journal of Remote Sensing, 25, pp.2365−2407.

Lund, H.G. 2009. What is a degraded forest? Forest Information Services. Gainesville, VA, USA.

Mackey, B.G., Keith, H., Berry, S., Lindenmayer, D.B., 2008. Green carbon: the role of natural forests in carbon storage. Part 1, A green carbon account of Australia's southeastern Eucalypt forest, and policy implications. ANU E Press, Canberra, Australia.

MacLean, I.M.D., Suggitt, A.J., Wilson, R.J., Duffy, J.P., and Bennie, J.J. 2017. Fine-scale climate change: Modelling spatial variation in biologically meaningful rates of warming. Global Change Biology, 23 pp.256–268.

Masek, J.G., Vermote, E.F., Saleous, N., Wolfe, R., Hall, F.G., Huemmrich, F., Gao, F., Kutler, K., and Lim, T.K. 2013. LEDAPS Calibration, Reflectance, Atmospheric Correction Preprocessing Code, Version 2. ORNL DAAC, Oak Ridge, Tennessee, USA.

Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., and Lim, T.-K. 2006. A Landsat surface reflectance dataset for North America, 1990–2000. IEEE Geosci. Remote Sens. Lett., 3, pp.68–72.

Maurer, E.P., Hidalgo, H.G., Das, T., Dettinger, M.D., and Cayan, D.R. 2010. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrol. Earth Syst. Sci., 14, pp.1125–1138,

McDermid, J., Fera, S., and Hogg, A.A. 2015. Climate change projections for Ontario: an updated synthesis for policymakers and planners. Ontario Ministry of Natural Resources and Forestry, Science and Research Branch, Peterborough, ON. Climate Change Research Report CCRR-44.

McKenney, D.W., Pedlar, J.H., Lawrence, K., Gray, P.A., Colombo, S.J., and Crins, W.J. 2010. Current and projected future climatic conditions for ecoregions and selected natural heritage areas in Ontario. Ontario Ministry of Natural Resources, Applied Research and Development Branch, Sault Ste. Marie, ON. Climate change research report CCRR-16.

Nabuurs, G.J., Masera, O., Andrasko, K., Benitez-Ponce, P., Boer, R., Dutschke, M., Elsiddig, E., Ford-Robertson, J., Frumhoff, P., Karjalainen, T., Krankina, O., Kurz, W.A., Matsumoto, M., Oyhantcabal, W., Ravindranath, N.H., Sanz Sanchez, M.J., and Zhang, X. 2007. Forestry. In, B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (Eds.), Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Nadeau, C.P., Urban, M.C., and Bridle, J.R. 2017. Coarse climate change projections for species living in a fine-scaled world. Global Change Biology, 23, pp.12–24.

NASA. MODIS gap-filled yearly NPP products (MOD17A3HGF-V061). modis.gsfc.nasa.gov.

Ontario Ministry of Natural Resources. 2008b. The boreal forest. Ontario Ministry of Natural Resources.

Perera, A.H., and Euler, D.L. 2000. Landscape ecology in forest management: An introduction. In, A.H. Perera, D.L. Euler, and I.D. Thompson (Eds.), Ecology of a managed landscape: patterns and processes of forest landscapes in Ontario, (pp.3–11). University of British Columbia Press, Vancouver, Canada.

Qiu, S., He, B., Zhu, Z., Liao, Z., and Quan, X. 2017. Improving Fmask cloud and cloud shadow detection in mountainous area for Landsats 4–8 images. Remote Sens. Environ., 199, pp.107–119.

Qiu, S., Lin, Y., Shang, R., Zhang, J., Ma, L., Zhu, Z., Qiu, S., Lin, Y., Shang, R., Zhang, J., Ma, L., and Zhu, Z. 2019. Making Landsat time series consistent: evaluating and improving Landsat analysis ready data. Remote Sens., 11(1), pp.51.

Qiu, S., Zhu, Z., and He, B. 2019. Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery. Remote Sensing of Environment (in review 2018; published 2019).

Romero-Sanchez, M.E., and Ponce-Hernandez, R. 2015. The assessment of forest degradation in dry forested lands: Mapping regional trend indicators of degradation in the Yucatan Peninsula, Mexico with satellite data. In, Hubert, B., and Broin, M. (Eds.), 3rd UNCCD Scientific Conference: "Combating desertification/land degradation and drought for poverty reduction and sustainable development: The contribution of science, technology, traditional knowledge and practices." Agropolis International, Montpelier, France, pp. 284–285.

Running, S.W., and Zhao, M.S. 2015. User's Guide. Daily GPP and Annual NPP (MOD17A2/A3) Products NASA Earth Observing System MODIS Land Algorithm. Version 3.0 for Collection 6. https://lpdaac.usgs.gov/sites/default/ files/public/product_documentation/mod17_user_guide.pdf.

Schindler, D.W. 1998. A dim future for Boreal waters and landscapes. Bioscience, 48, pp.157–164.

Schmidt, G.L., Jenkerson, C.B., Masek, J., Vermote, E., and Gao, F. 2013. Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description: U.S. Geological Survey Open-File Report 2013–1057.

Schroeder, T.A., Cohen, W.B., Song, C.H., Canty, M.J., and Yang, Z.Q. 2006. Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon. Remote Sensing of Environment, 103(1), pp. 16-26.

Scott, D., and Lemieux, C. 2005. Climate change and protected area policy and planning in Canada. Forestry Chronicle, 81(5), pp.696–703.

Seidl, R., Schelhaas, M.J., Rammer, W., and Verkerk, P.J. 2014. Increasing forest disturbances in Europe and their impact on carbon storage. Nature Climate Change, 4, pp.930–930.

Singh, A. 1989. Digital change detection techniques using remotely-sensed data. International Journal of Remote Sensing, 10, pp.989–1003.

Soja, A.J., Tchebakova, N.M., French, N.H.F., Flannigan, M.D., Shugart, H.H., Stocks, B.J., Sukhinin, A.I., Parfenova, E.I., Chapin, F.S. III, and Stackhouse, P.W. Jr. 2007. Climate-induced boreal forest change: predictions versus current observations. Global Planet Change, 56, pp.274–296.

Stern, N. 2007. The economics of climate change: The Stern review. Cambridge, Cambridge University Press.

USGS. EarthExplorer; https://earthexplorer.usgs.gov

USGS. Global Visualization (GloVis); https://glovis.usgs.gov

USGS. 2016. Product Guide. Provisional Landsat 8 Surface Reflectance Code (LASRC) Product. ‘‘Department of Interior, U.S. Geological Survey” <http://landsat.usgs.gov/CDR_LSR.php>.

USGS. 2017. Landsat 4–7 Climate Data Record (CDR) Surface Reflectance. <http://landsat.usgs.gov/documents/cdr_sr_product_guide.pdf>.

USGS. 2019. Landsat 8 (L8) Data Users Handbook Version 5.0. Department of Interior, U.S. Geological Survey.

USGS. 2020. Landsat 4–7 Collection 1 (C1), Surface Reflectance (LEDAPS) Product Guide, Version 3.0. Department of Interior, U.S. Geological Survey.

USGS. 2020. Landsat 8 Collection 1 (C1), Land Surface Reflectance Code (LaSRC) Product Guide, Version 3.0. Department of Interior, U.S. Geological Survey.

Vermote, E., Justice, C., Claverie, M., and Franch, B. 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. Remote Sensing of Environment, 185, pp.46–56.

Wulder, M.A., Campbell, C., White, J.C., Flannigan, M.D., and Campbell, I.D. 2007. National circumstances in the international circumboreal community. For Chron, 83(4), pp.539–556.

Zhu, Z., and Woodcock, C.E. 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens. Environ., 118, pp.83–94.

Zhu, Z., Woodcock, C.E., and Olofsson, P. 2012. Continuous monitoring of forest disturbance using all available Landsat imagery. Remote Sens. Environ., 122, pp.75–91.

CHAPTER TWO

LAND COVER CLASSIFICATION

Evaluation of the accuracy of multispectral land cover classification algorithms: Comparison of Machine Learning, Maximum Likelihood, and Mahalanobis Distance Algorithms

Mozammel Md. Hoque[a], R. Ponce-Hernandez[b], C. Gibson[c], T. Jones[d]

[a]Md Mozammel Hoque (submitting and corresponding author)
PhD Candidate, Environmental and Life Sciences
Trent University
1600 West Bank Dr, Peterborough, ON K9L 0G2
Present address: 700 Sawmill Road, Peterborough, ON K9L 0E8
Tel: (647) 208-0384
Email: mdhoque@trentu.ca

[b]Dr. Raul Ponce-Hernandez
Professor, School of the Environment
Trent University
1600 West Bank Dr, Peterborough, ON K9L 0G2
Tel: (705) 748-1011 ext. 7646
Email: rponce@trentu.ca

[c]Dr. Carey Gibson
(Ex-Remote Sensing Specialist, OMNRF, Peterborough)
Present address: 1284 Cartier Blvd, Peterborough, ON K9H 6S1
Tel: (705) 749-2084
Email: carey.gibson@sympatico.ca

[d]Dr. Trevor Jones
Research Scientist
Canadian Wood Fibre Centre
Natural Resources Canada, Government of Canada
1219 Queen St., East, Sault Ste. Marie, ON P6A 2E5
Tel: (705) 989-4352
Email: trevor.jones@nrcan-rncan.gc.ca

Evaluation of the accuracy of multispectral land cover classification algorithms: Comparison of Machine Learning, Maximum Likelihood, and Mahalanobis Distance Algorithms

Hoque, Md Mozammel, Raul Ponce-Hernandez, Carey Gibson, Trevor Jones

Abstract

The objectives of this paper are to examine and evaluate the accuracy of conventional and relatively new algorithms for land cover classification based on Landsat imagery of Central and Northeastern Ontario, Canada. To ascertain the best supervised classifier, we compared the accuracy of resulting land cover maps produced by four algorithms: two non-parametric machine learning algorithms: support vector machine (SVM) and artificial neural networks (ANN), as well as two parametric classification algorithms: maximum likelihood classifier (MLC) and Mahalanobis distance (MahD).

The overall accuracy values obtained by SVM, MLC, MahD, and ANN were 98.6%, 85.7%, 83.6%, and 76.8%, respectively. Additionally, the results showed MLC, MahD, and ANN classification algorithms are highly influenced by imbalanced training data that can lead to lower accuracy, while SVM is not affected. The findings revealed that SVM is the most robust, consistent, and effective classifier of those compared. While a greater amount of time and effort was spent in making improvements to the signature file to increase the accuracy from 60.86% (14 hrs) to 98.62% (42 hrs) with the SVM algorithm, more research is needed to ascertain whether the gains were truly proportional to the extra time, effort, and expense.

Keywords: Remote sensing, Land cover, Image classification, Imbalanced training dataset, F1 score, Tuning parameters, SVM, ANN, MLC, MahD.

2.1 Introduction

Land cover/land use (LCLU) mapping is a subject of great interest due to its multiple applications (Soffianian and Madanian, 2015; Wang et al., 2015) in planning and modeling ecosystem activities (Shaw et al., 2014; Bhattarai et al., 2015; Talukder et al., 2020). Understanding LCLU is important in habitat

distribution, biodiversity, global change studies (Verburg et al., 2011; Hansen, T., 2012), LU policy development, urban planning, nature conservation, agricultural monitoring, as well as studying the impacts of climate change (Shetaii and Abdi, 2008; Talukder et al., 2020). Therefore, there is a need to create accurate and current LCLU maps. Along with conventional field methods, remote sensing (RS) techniques applied to satellite imagery furnish opportunities to monitor and assess LCLU types. The identification and mapping of LCLU types through multispectral satellite image classification are key to the usefulness of the products. However, it is still a challenge to obtain reliable and accurate LCLU classifications from RS images due to a number of factors, such as landscape complexity, RS data selection, image processing, and particularly the choice of classification algorithms. Image classification can be a lengthy and laborious process and usually prone to inaccuracies. Developing an effective image classification with acceptable levels of accuracy demands time and effort, and it is dependent on the effectiveness of the algorithms used for the generation of a classification map.

Although relatively new machine learning algorithms for image classification have been developed in the past few years, only a handful of studies have compared their applicability and performance. As information on forest composition is crucial for forest management and inventory, RS data and image processing techniques have been widely employed in mapping the spatial distribution of forest types, at various spatial and temporal scales (Thessler et al., 2008; Aguirre-Salado et al., 2012). While accuracy depends on a variety of factors, including the targets requiring classification and the inputs employed, RS, with its increased spatial and spectral resolution, is both accurate and much easier than ground survey (although ground survey is more accurate than RS). Among the synoptic and periodic observations of Earth Observation (EO), Landsat images are widely employed for mapping the Earth's surface (Williams et al., 2006). Since the advent of Landsat 1 in 1972, Landsat images have often been employed for the purpose of forest monitoring due to their suitable spectral and spatial resolutions (Huang et al., 2008; Roy

et al., 2014), and much research has focused on forest time series analysis based on the large time series of satellite image stacks. Analyzing such multi-spectral images can aid in LC identification (Tigges et al., 2013; Shim, 2014). In particular, RS techniques, when used wisely and effectively, offer an approach that is accurate and time-saving when obtaining LCLU information.

The image classification process is a multi-step workflow that involves extracting information from satellite imagery by separating the image pixel values into meaningful categories or LC classes (Chaichoke et al., 2011). Scene classification of identifiable or meaningful features or LC type classes results in a map such as LU and types of vegetation (Palaniswami et al., 2006). LCLU mapping is a subject of great interest due to its multiple applications (Soffianian and Madanian, 2015; Wang et al., 2015). Understanding LCLU is important in habitat distribution, forest management and inventory, evaluating aquatic ecosystems, biodiversity, global change studies (Verburg et al., 2011; Hansen, T., 2012), LU policy development, urban planning, conservation and environmental monitoring, agricultural monitoring, as well as studying climate change (Sellers et al., 1995; Shataeii et al., 2008). Therefore, there is a need to create accurate and current LCLU maps.

Despite the advantages of RS detailed above, creating an accurate map through the classification of RS imagery does present some challenges due to many factors, including landscape complexity, the choice of RS data, as well as image processing and classification approaches that can affect classification accuracy (Lu and Weng, 2007). Of particular importance is the selection and evaluation of a suitable classification algorithm along with a required number of "training" samples from field sampling and the selection of data sources (e.g., aerial photographs and satellite imagery).

Pixel-based image classification is generally employed for information extraction and data analysis. The classification accuracy is affected by the resolution, purity, and size of the training dataset, heterogeneity of LC class as manifested on the ground, and the type of adopted classifier (Lu and Weng, 2007; Kavzoglu,

2009). The quality and characteristics of training data are key for classification accuracy. Additionally, when classifying the minor classes of interest, the size of the training dataset acquires significance. In some cases, there could be limitations when collecting the sampled pixels (and ground truth samples) for a specific class. This results in an imbalanced training dataset as the number and size of training samples of one specific class is smaller than those of the other classes. While this may lead to low accuracy, some classifiers could be highly accurate even under those conditions (Foody et al., 2006; Waske et al., 2009). Some studies have examined how to deal with the problem of classification with imbalanced data by resampling the training dataset and using classifiers such as support vector machine (SVM) (Waske et al., 2009; Mohd Pozi et al., 2015).

According to Lu and Weng (2007), both imagery appropriateness and the correct choice of classification algorithms affect the performance of LCLU mapping results. Numerous image classification algorithms have been developed since the first Landsat image was acquired in the 1970s (Bakr et al., 2010; Yang et al., 2011). Classification methods for LCLU mapping using RS data range from unsupervised algorithms (i.e., ISODATA or K-means) to parametric supervised algorithms, such as maximum likelihood (ML) or Mahalanobis distance (MahD), and nonparametric machine learning supervised algorithms. Among the latter, artificial neural networks (ANN), SVM, decision trees (DT), k-nearest neighbors (kNN), and random forest (RF) are often used.

Much research has compared different image classification algorithms for LCLU mapping (Dixon and Candade, 2008; Khatami et al., 2012; Talukder et al., 2020). However, conclusions vary widely. Therefore, it is necessary to compare the different image classification algorithms in different environments, such as, in this case, the temperate and humid area in Ontario, Canada, selected for this study. For this area, the LC classification of Landsat data was performed using four supervised classifiers: two parametric algorithms (ML and MahD) and two non-parametric, machine learning algorithms (ANN

and SVM). The performance of these four algorithms was compared in terms of the classification accuracy obtained. The learning stage of the supervised classification algorithm is mainly dependent on the training data characteristics, specifically on the selection of the training sub-populations and their location on the input imagery, as well as the implemented classificatory algorithms. The sample pixel distributions are also vital in terms of classification performance, especially for parametric classifiers (e.g., ML) (Lu and Weng, 2007; Kavzoglu, 2009). The imagery and training sets are also crucial for obtaining high accuracy in the classification, as they rely on supervised learning (Kavzoglu, 2009). In order to successfully assess the influence of imbalanced training datasets on the accuracy of image classification, four different pixel-based supervised classification methods (SVM, ML, MahD, and ANN) were selected and implemented for mapping LC types, and then they were compared in terms of the resulting accuracy of each classificatory algorithm. Performance of the classifiers depends on many factors (e.g., selection of training samples, the data samples selected for testing, and the input variables) (Foody et al., 1995).

Over the last few decades, ANN classification has been extensively employed in RS as an alternative to statistical classification methods, a highly self-learning and self-adaptive algorithm with the capability to recognize and generalize complex datasets (Yu and Chen, 2020; Talukder et al., 2020; Ghayour et al., 2021). ANN has been argued to be better than traditional statistical classification approaches (such as ML and MahD) because it does not make assumptions about the nature of the probability distribution function, which is learned from training samples (Foody, 2004). This back-propagation algorithm has been used extensively for a wide variety of applications (Mas and Flores, 2008), and it is often chosen specifically due to the fact that it is a non-parametric and non-linear classifier. However, ANNs can be problematic as they need a lot of computational power to obtain the desired output, are “slow learners,” and require huge training datasets.

SVM, which is based on statistical learning, is a supervised non-parametric machine learning algorithm that can be employed for the creation of LCLU maps (Jensen, 2005; Talukder et al., 2020). Over the last 20 years, SVM and ANN algorithms have been popular for image classification (Pal et al., 2013). Generally, SVM is employed due to its potential for better classification accuracy than ANN (Foody and Mathur, 2004; Pal et al., 2013). SVM's main advantage is its ability for generalization with limited training samples. While SVM requires parameter selection and some computational demands, in certain instances it has been shown to offer superior performance compared to other image classification algorithms, and it could furnish good classification results from data that are complex and noisy. In order to ensure there is little chance of misclassifying previously unseen data, the algorithm makes use of a structural risk minimization principle (Vapnik, 1998). Overall, the SVM classifier could offer an innovative approach compared to traditional parametric classifiers (e.g., ML and MahD), as well as to the non-parametric ANN classifier (Mountrakis et al., 2011; Srivastava et al., 2012).

Finally, there is an immense amount of time and effort that must be invested to ensure classification accuracy; however, this problem has been underrepresented in the literature, in particular regarding "signature/reference" file preparation and achieving accuracy values. Thus, in this research, metrics of time and effort were developed and used in the classification process, and they were evaluated as part of the "cost" of obtaining the performance of the algorithms in terms of the accuracy of the results.

This paper examined and evaluated a selected group of conventional and relatively new advanced methods and techniques for multispectral satellite image classification (SVM, MLC, MahD, and ANN) in terms of their accuracy in LC classification of Landsat imagery for central and northeastern Ontario, Canada. The creation of a map of the study area with the highest possible level of accuracy was intended, since no map of this kind exists in the area. To support and validate the classification assessment, ground truth points were used to ascertain the OA, K coefficient, and F1 score metrics. A central goal of this paper was to

achieve a classification accuracy beyond 90%, and to examine the "cost" in terms of time and effort necessary for this achievement. After identifying the algorithm that produced the highest classification accuracy, additional investigations examined the relationships between classification accuracy and the "cost" in terms of time and effort required to attain it by each algorithm. To this end, the field was revisited multiple times for the collection of more ground truth data to review potentially misclassified factor (LC type) areas until reaching the highest accuracy level possible. Finally, the influence of imbalanced training data on the resulting classes of LC types was evaluated.

2.2 Materials and Methods

2.2.1 Study Area

The area selected for this study is part of the Algonquin Park area in Ontario, Canada. Its spatial extent, spans from 45°8'15.32''N to 46°16'12.38''N and from 76°37'48.75''W to 78°20'16.73''W, defining a total of 10,000 km^2. The perimeter of the area was digitized based on the Electoral District of Renfrew-Nipissing-Pembroke (ED088) map. ED088 includes all of Renfrew County and a small section of the Nipissing District, which is around Algonquin Provincial Park. The original ED088 district area comprising the study area was 12,383.83 km^2. The final study area, after exclusions of water areas of the

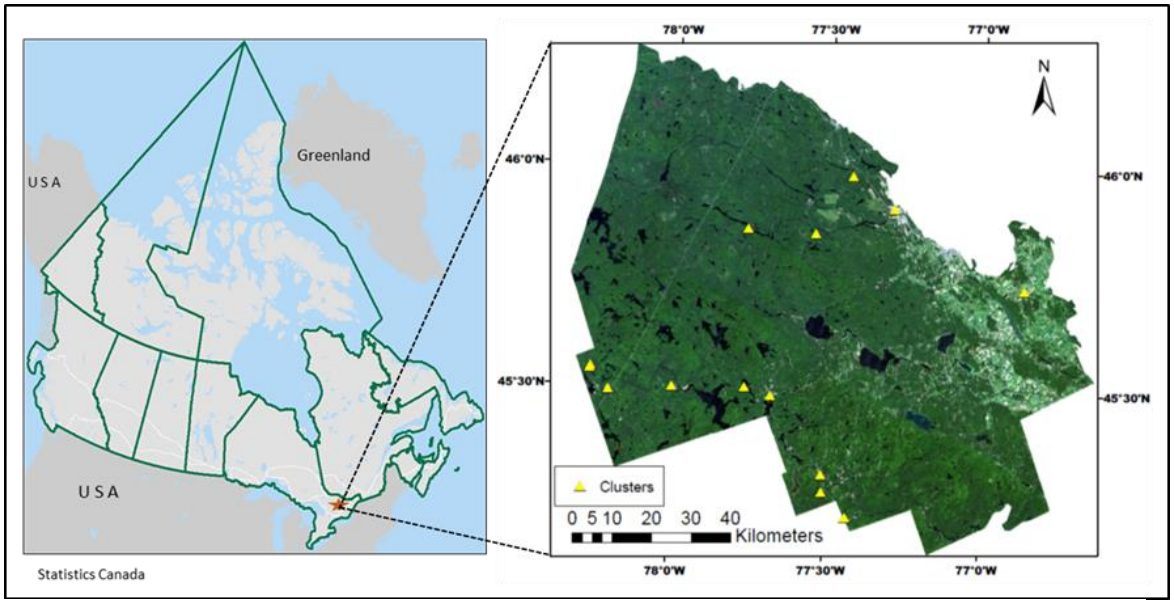


Figure 2. 1: Study area with clusters in central and northeastern Ontario, Canada

Ottawa river and map corners was 10,000.02 km^2. The Lambert Conformal Conic projection was converted to WGS 1984, UTM Zone 18N to keep the identical spatial reference of the acquired Landsat images. This study area (Figure 2.1) contains a wide range of forest and non-forest LC types, as it is typical of forests in the boreal transitional zone, including Algonquin Park.

The study area is typically characterized as two distinct forested regions, namely North East Algonquin (Nipissing District and Renfrew County) and South West Algonquin (Bancroft, Minden and French Severn Forests). The southwestern portion is in a higher elevation and has sandy loam soils derived from glacial till, which tends to have high moisture holding capacity and is dominated by deciduous forest types. The northeastern side consists of coarse outwash soils and is dominated by white pine and mixed deciduous forests. The main tree species include red oak, red and sugar maple, trembling and large-tooth aspen, basswood, ironwood, white and yellow birch, beech, black and white ash, among others, and the coniferous tree species include mainly red, white, and jack pine; black, white, and red spruce; balsam fir; hemlock; and tamarack, among others (Chubey et al., 2009).

The study area generally has a humid continental climate (Köppen, 1936) with long, snowy, and cold winters and warm, moist summers. The average annual temperature and precipitation of Algonquin Park during the period of 1971–2000 was 4.5°C and 949.0 mm, respectively (McKenny et al., 2010; Favot et al., 2019). Of this precipitation, approximately 25% is snowfall. On average, there are 136 days in the growing season, with 100 frost-free days (Chubey et al., 2009). In our study, for the periods from 1971–2000, 1981–2010, and 2011–2020, we found the average temperature and precipitation for the study area were 4.9°C and 1022.0 mm; 5.1°C and 1085.5 mm; and 5.2°C and 949.0 mm, respectively. However, there were some missing precipitation data from most of the stations during the period from 2011–2020.

2.2.2 Field Sampling

The initial study area was 100 km^2 in the cluster of the Petawawa Research Forest, where most of the samples were collected due to the accessibility. Subsequently, the area was expanded to encompass 10,000 km^2, making field visits difficult and time consuming. Accessibility was an important and determining factor strongly influencing data collection and the spatial distribution of samples. A randomized cluster sampling design was adopted for this heterogeneous study area, as it was difficult to collect field data due to the limitations of access roads and the risk posed by potentially dangerous wild animals inside the forest, meaning that accessibility was a factor influencing data collection, especially in terms of the number and spatial distribution of samples. So, the time, effort, and expense involved in collection resulted in a limited, and sometimes unbalanced, number of ground samples. In total, 14 clusters of ground samples were part of the sampling effort to provide training sites for the classification efforts. Ground measurements were collected from 250 locations within each of the 14 sampling clusters in the study area (Figure 2.1). To fix the misclassified LC types, sample re-collection was undertaken; however, these were in addition to, not included in, the 250 samples. A circular plot of 50m in diameter and containing 50m transects was established at each of the locations (Figure 2.2), and one location contained 31 measurements for tree data (e.g., species, dominance, percentage of canopy cover, etc.). The number of transects depended on the local conditions of access. LC types were identified visually and recorded. Other variables, such as the height of trees and their diameter at breast height, were also measured and recorded, even though these variables were collected for parallel studies and are only supporting, not determinant, variables used in the classifications. All data were collected between July and September 2018 (the area's growing seasons).

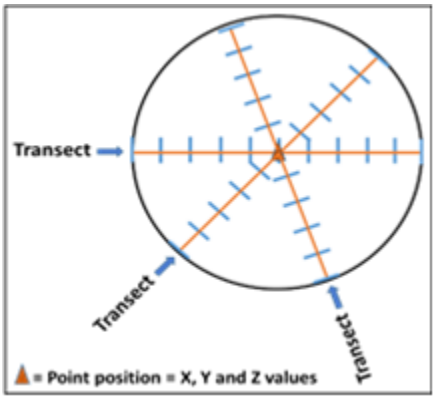


Figure 2. 2: Collection of point data from the field using transect

The field data collection identified and recorded all LC types in and around each cluster/site. The sampling procedure undertaken on the ground initially identified 14 LC types (in version 1), which were subsequently refined into 7 LC types (in version 8) (Table 2.1, below). This field identification of LC types on the ground was implemented based on the study objectives and careful consideration of an initial analysis of the separability of LC types.

The sampled locations showed dominance (with more than 80% canopy cover) of conifer and deciduous forest LC types. Mixed forest samples showed an almost equal presence of deciduous (50%) and coniferous (50%) forests. The class "water" was represented throughout the study area with the generic class "wetlands," which were grouped into a few samples of the class "bogs," "marshes," and "swamps." The sampling also included a few samples of the class "infrastructure." The latter were identified using finer resolution satellite images. All LC type data were gathered from the ground and located using a handheld GPS with near meter accuracy, collecting data from clusters 1 to 13. The sampling clusters were located on the satellite image and searched for and found on the ground within an accuracy of 2m to 7m. A simple random sampling technique was applied to locate the sampling clusters and data points, of which 70% were used for training and 30% for testing the classification results. Based on accessibility, the number of samples for each individual LC type ranged from 15 to 73 samples (Table 2.1).

2.2.3 Methodological Workflow

The overall methodology for this project consisted of a series of analytical steps that, when put in sequence, made up a methodological workflow, which could be organized as a flowchart. The workflow for the methodology in this project is shown in Figure 2.3. The first block of activities consisted of steps related to the preparation of an image stack and image correction processes, consisting of the SLC gap filling and cloud/cloud shadow masking procedures. The resulting corrected image stack allowed for the selection of the best Landsat image for the year 2018, which was used for the LC classification objectives.

The last step of the first block of procedures involved the selection of the "best" classification algorithms for their comparison in terms of classification accuracy and performance components. The procedures for accuracy assessment using ground validation samples allowed for the identification of the "best" algorithm in terms of accuracy of the classification.

The second block of activities consisted of procedures for examining the time and effort investment. To increase the classification accuracy, eight versions (V1–V8) of signature files (ROIs/reference files) were created to train the algorithm to obtain the highest level of acceptable classification accuracy.

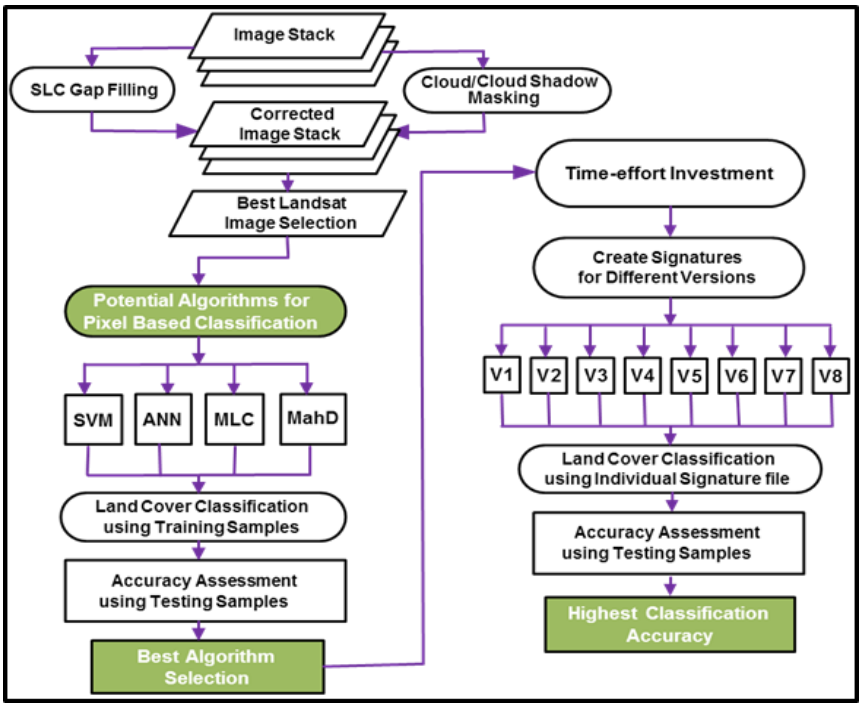


Figure 2. 3: Flow diagram for the best algorithm selection and time-effort investment for the highest accuracy. Note that V1 to V8 represent the number of versions of the classification required to achieve the highest level of accuracy.

2.2.4 Data

2.2.4.1 Landsat Imagery and Preprocessing

Landsat data were chosen in particular because of their availability for long-term assessment in the study area (from 1972–2020), which provided a more detailed picture for future projections. (Data from other platforms, such as SAR and Sentinel, are not as readily available in comparison with Landsat imagery). Multiple Landsat images were acquired during leaf-on seasons, from June 21 to September 22 (summer solstice to fall equinox), which coincides with the peak of the growing season, to ensure that the images had less than or equal to 10% cloud cover. Landsat image preprocessing included the detection and restoration of bad lines, radiometric correction, gap filling, cloud and cloud shadow removal, as well as

composite filtering. Landsat 1–5 MSS, Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat 8 Operational Land Imager (OLI) were collected for the period from 1972–2020 for this study from USGS platforms (e.g., glovis.usgs.gov). Digital Numbers were not converted to reflectance. Landsat TM and ETM+ data were acquired for the period from 1984–2013, as well as collected Surface Reflected data (collection-1, level-2), using the USGS "On Demand" option. Landsat 8 data for 2013–2020 were also acquired, and Surface Reflected data (collection-1, level-2) were collected using the same technique. A cloud-free Landsat 8 image of September 2018 was selected for LC classification after preprocessing and correcting the image stack. The study area was covered by one scene/tile of path 017 (MSS 018) and row 028 (no mosaic). Landsat 8 OLI surface reflectance products were obtained already corrected by the LaSRC algorithm (USGS, 2016, 2020). These corrections increased the radiometric resolution and improved signal-to-noise ratio (Vermote et al., 2016; USGS, 2019). As a result, Landsat 8 surface reflectance product is far superior to the makeshift Landsat 5/7 LEDAPS product (Vermote et al., 2016; USGS, 2019). Fmask 4.0 (beta version), which automatically detects water, snow, clouds, and cloud shadows, was employed for Landsat 4–8 and Sentinel-2 images, which is developed based on Fmask and MFmask algorithms version 3.3 (Qui et al., 2017, 2018, 2019). Landsat 1–5 MSS images were corrected with the abovementioned approaches using Exelis Visual Information Solutions (ENVI ^{TM} 5.3). As it is not possible to calibrate or atmospherically correct old MSS data in a reliable manner, every scene from the 1975-era MSS (Landsat 1–5 MSS) archive was fixed radiometrically in agreement with the ETM+ image (as in Hall et al., 1991). For more details about Landsat imagery and preprocessing, see Chapter 1.

2.2.4.2 Sample Datasets for Training and Testing the Classifiers

Since a large number of published works (e.g., Stehman, 1992) have already examined the impact of the choice of testing data on classification accuracy, in this research, emphasis was placed on the selection of training samples and input variables. The training dataset should have a sufficient number of independent

sample units of LC types for each LC class to exhibit interclass variability (Muñoz-Marí et al., 2007; Ghayour et al., 2021). Based on statistical theory, we applied specifically categorized training areas to recognized similar areas for each individual LC class. In many studies, the sample datasets are divided into the categories of training and testing. To ensure a high level of accuracy for the classifiers, the recommended proportion is using 70% of the data to train the classifiers and 30% for validation (Yousefi et al., 2015; Pouliot et al., 2019).

To ensure the independence of the training and testing sample datasets for classification, a simple random sampling technique was implemented to select the training and testing sample datasets for calibration (subsequently referred to as the calibration set). The calibration datasets were kept the same for all classifiers to avoid sampling bias. Therefore, to train the four classifiers, 175 (70%) samples were randomly selected for training and the remaining 75 (30%) were used for testing to evaluate the performance of the classifiers. To address the misclassified LC types, the site was revisited to collect new samples; however, these were not included in the 250 samples.

2.2.4.3 Training and Testing Signatures

The training signature file (i.e., reference or "ROI" file in the ENVI^{TM} software) for image classification can be created using fine resolution satellite imagery and field sample data with balanced and imbalanced samples. As indicated earlier, a total of 250 imbalanced samples were collected from the field (not counting the additional samples to address the misclassification mentioned above). Of these, 175 were randomly selected to identify areas of the Landsat 8 image from which to collect training signatures. To obtain a representative spectral signature of the class, individual regions of interest (ROIs) were drawn to include essential pixels. The remaining 75 ground truth samples were employed in the accuracy assessment to compare the performance of the four classifiers. A signature file was created for training the four classifiers using ENVI^{TM} software with ground truth/reference data collected from the field. It

should be noted that for SVM, there were 8 signature files created, one for each of the 8 versions. Another signature file was created using 75 testing samples to compare the performance of the four classifiers and select the best classifier. In order to test the performance of SVM, a total of 8 signature files were also created, one for each of the 8 versions.

2.2.4.4 Filtering

Individual and combined filtering were used for this study. A combination of Median (Haralick et al., 1987) and Local Sigma filters (Eliason and McEwen, 1990) were employed. Combined filters increased the smoothness and brightness of the image, as well as removed all kinds of noise, such as spot noise, speckle noise, etc. All classificatory algorithms were applied after these image enhancement procedures. However, it was noted that the SVM classification results increased in accuracy by 3.28% after implementation of the combined filters.

2.2.4.5 Classification

Supervised classification labels pixels according to user-defined classes based on previous knowledge of the features to be categorized. The training and testing/validation datasets were common for each algorithm. Further, owing to the ease of selection and the concomitant decrease in misclassification error, pixels in the study area were allocated into the seven dominant LC classes, which are shown in Table 2.1.

Table 2. 1: Classified LC classes in the study area including the training and testing sample units

LC classes	**Training samples**	**Testing samples**	**Description of LC types**
Coniferous	30	13	Coniferous and shelterwoods forest
Deciduous	15	7	Deciduous forest
Mixed Forest	18	8	Mixed forest
Grassland	10	5	Crops, pastures
Wetland	40	16	Bog, marsh, swamp, dead/dry trees
Infrastructure	52	21	Buildings, roads, asphalt, parking lots (gravel and sandy soil)
Water	10	5	Lakes, rivers, and reservoirs

There were some minor imbalances in terms of the different classes in the training and testing samples, for which it should be noted that several classes, such as wetland and infrastructure, are composed of a

number of types of LC. However, the imbalances in the testing samples were mitigated by merging categories in SVM from V1–V8. While in Table 2.1, it can be seen that there are higher numbers of testing samples of wetland (16) and infrastructure (21) (which are equivalent to 21% and 28% of the total testing samples), this was only for the first version. In V8, the testing samples for wetland (8) and infrastructure (5) were much more balanced and representative of the actual distribution of classes in the field.

2.2.4.6 Tuning Parameters

Tuning parameters for classification, as used by the ENVI 5.3 software, have a significant part to play in ensuring highly accurate results of image classification with SVM, ANN, ML, and MahD. All four classifiers have different tuning parameters and procedural steps. This study tested a series of parameter values for each classifier's tuning process, with the optimal parameters identified based on the highest overall accuracy of the classification results (Than Noi and Kappas, 2017). Such optimal parameters were subsequently employed in this study, when comparing the efficacy of the various classifiers.

Accuracy metrics were used to assess the quality and performance of alternative classifiers. Therefore, in this study, each classifier was run separately, iterating through all parameter grid points and input datasets, after building the spectral signature file (reference file or ROI file in ENVI software), and using training datasets, setting up the tuning parameters. The results of the tuning process used a performance metric to rank the classifiers according to different parameters, in order to optimize a parameter for each classification algorithm.

In this study, a confusion matrix and accuracy metric were employed to measure classification performance (with testing samples) and conduct a search of the values to determine the optimal parameter with which to train the classifier. The ideal parameter values of the four classifiers were found by implementing all combinations of the aforementioned parameter values for this study. Following the studies of Than Noi and Kappas (2017) and Li et al. (2014), 10 iterations were tested on ranges of values

for the kernel coefficient (γ) and cost (C), to find the optimal parameters of SVM. Optimal values found for the parameter setting are shown in Table 2.2, in order to reach the highest overall accuracy (OA).

Table 2. 2: Tuned parameters for each supervised classifier and tuned optimum parameter value used

Supervised classifier	Tuned parameters	Search range	Optimal parameter Setting
SVM	Kernel coefficient for radial basis function (RBF) (γ) Penalty parameter for error term cost (C) Classification probability threshold	0.0001–0.2 0.1–1000 0-1	0.143 100 0
ANN	Training threshold contribution and momentum Training/learning rate Maximum number of learning iterations	0.5–0.9 0.1, 0.2, 0.5 200–1000	0.6 0.2 1000
MLC	Probability threshold = no threshold, single value, or multiple values	Single value from 0–1	Single value
MahD	Maximum distance error value	None, single value, or multiple values	Single value

2.2.4.7 Classification Algorithms

2.2.4.7.1 Support Vector Machine (SVM)

SVM is being applied more frequently for LC classification owing to its ability to solve classification problems with a small number of training samples and its nonlinearity (Brown et al., 1999). It is a high-performing supervised classification algorithm with four core models: linear, polynomial, radial basis function (RBF), and sigmoid kernels (Brereton and Lioyd, 2010). In their study, Camps-Valls et al. (2008) presented a new group of kernel-based methods for time series image classification. SVM employed kernel functions to divide classes in a higher dimensional space if two classes were not able to be separated linearly. Kernel functions require user-defined parameters and the option of a suitable kernel type as the corresponding parameters have a significant impact on SVM performance as it helps in the identification of the "shattering points" or "support vectors" (Melgani and Bruzzone, 2004; Talukder et al., 2020). In general, RBF kernel generates superior classification results (Kavzolgu and Colkesen, 2009; Otukei and Blaschke, 2010), meaning the SVM classifier's RBF kernel is commonly used, needs less computational work, and displays good classificatory performance (Knorn et al., 2009; Shi and Yang, 2015). Importantly,

the RBF kernel is generally used when there is a non-linear relationship between the training samples and the entire dataset. This means it is an effective classifier for heterogeneous areas (Su et al., 2007). When using the SVM classifier with the RBF kernel, two parameters need to be set: the optimum penalty parameter of cost (C) and the kernel width parameter (γ) (Qian et al., 2015; Ballanti et al., 2016). The shape smoothing of the class-dividing hyperplane is affected by the parameter γ (Melgani and Bruzzone, 2004). Larger C values may lead to an over-fitting model (Ghosh and Joshi, 2014), while increasing the γ value will influence the shape of the class-dividing hyperplane. This, in turn, may affect the accuracy of the classification results (Huang et al., 2002; Ghosh and Joshi, 2014). In this study, the RBF kernel employed for the implementation of the SVM algorithm included γ in all, the kernel function, the probability threshold, and the penalty parameter. The kernel function and penalty parameters were optimized through comparison of the classification results after adjusting them continuously through an iterative process.

2.2.4.7.2 Artificial Neural Network (ANN)

A multilayered feed-forward ANN performs a non-linear classification. The ANN system employs standard back propagation for supervised (machine) learning, and a simple model consists of one input layer, minimum one hidden layer, and one output layer. ANN creates a method from the input layer for the prediction (Kalantar et al., 2018; Tan et al., 2021). In satellite image classification, the topology consists of a neuron in the input layer representing one of the input features (e.g., one satellite image band), and each neuron in the output layer corresponds to one class (Srivastava et al., 2012). In the learning stage, it is the neural network system design, as well as the optimization of training parameters (e.g., learning rate), that are particularly crucial to the accuracy of results (Kavzoglu and Mather, 2003; Kavzoglu, 2009; Talukder et al., 2020). The learning process involves adjusting the node weights to decrease the difference between the output node activation and the output. Next, the error is back

propagated through the network, and weight adjustment is done using the recursive method (Kar and Kelkar, 2013). It is possible to achieve the desired output by alternating the node weights.

The iteration gradient algorithm of back propagation generally minimizes the root mean square error (RMSE) between the true output of a feed forward ANN of multiple layers and the intended output. The iteration process is ended when the RMSE achieves the optimal level. It is important to remember that the training iteration number should not be too small or too large. In this study, the optimal parameter values of the ANN classifier were obtained by implementing different combinations. To identify suitable training protocols, various combinations of learning rates and momentum values were also tested.

2.2.4.7.3 Maximum Likelihood (ML)

ML is the most widely used classification algorithm in RS (Bailly et al., 2007; Kar and Kelkar, 2013). Therefore, this research employed it, considering the spectral characteristics of the satellite images and existing knowledge of the LC study area. The ML classifier applies the highest pixel probability to classify and estimate the probability of a pixel belonging to each class (Munyati, 2000; Jog and Dixit, 2016). To calculate the mean and variances of the training dataset classes to be implemented in the classification, probabilities are estimated, and changes are examined in the brightness values of each class. In order to classify an unknown pixel of a spectral response pattern, the Gaussian ML classifier also employs variance and co-variance of population distributions of classes. This classification relies on the probability density function connected with a particular signature. Pixels are allocated to a similar class on the basis of a comparison of the posterior probability that they should belong to each of the signatures under consideration (Vibhute and Gawali, 2013).

2.2.4.7.4 Mahalanobis Distance (MahD)

This algorithm assumes the bands' histograms have normal distributions and provides maximum accuracy (Kar and Kelkar, 2013). Based on the correlation between variables by which different patterns can be identified and analyzed, this classifier assesses the similarity of an unknown to a known sample set. Its

approach, which differs from Euclidian distance (Talukder et al., 2020), takes care of the correlation of the dataset and it is scale-invariant (Vibhute and Gawali, 2013). In this method, all pixels are classified to the closest (i.e., shortest distance) training data (i.e., the ROI class).

2.2.4.8 Accuracy Assessment and Performance Evaluation of the Classification Algorithms

Classification accuracy assessments were conducted on the testing dataset after the Landsat 8 (2018) image was classified via the training dataset. The classification accuracy was assessed by examining and understanding the source of the errors (Congalton and Green, 2019; Powell et al., 2004). The errors included classification and other types of errors, such as positional and interpretation errors, as well as poor quality training or testing samples. The error matrix technique was employed for accuracy assessment (Foody, 2002), as well as the overall accuracy (OA) and kappa coefficient (K) metrics. These were all used in this study to ascertain the accuracy of classification from each of the algorithms (Aguilar et al., 2012; Srivastava et al., 2012). K was employed since it is the most common assessment coefficient (Aguilar et al., 2012) as it makes use of pixels in the wrong classes (Smeeton, 1985) and describes the similarity between what is on the ground (ground truth/training datasets) and what is on the screen/paper (classification data) (Richard, 1999). K is calculated with the following equation:

$$\mathrm{K} = \frac{P_0 - P_C}{1 - P_C} \tag{2.1}$$

where, P_O = correctly observed pixels, P_C = hypothetical probability of the agreement of chance (expected agreement).

It is crucial to validate the quality of LC products for operational applications (Cihlar, 2000; Foody, 2002). Therefore, this process must involve appropriate sampling strategies for the statistical assessment of product accuracy. In addition, to ensure robust validation, the reference data must be of sufficient accuracy (Wulder et al., 2006). In particular, for multiclass models with imbalanced data, recall and precision are informative metrics (Buckland and Gey, 1994). While precision is a measure of exactness in terms of the

correct number of class predictions, recall is a measure of completeness in terms of how well a class is detected. It is also possible to talk about precision and recall in terms of error (commission or omission). Thus, for performance evaluation, frequently used accuracy measures (i.e., precision/sensitivity, recall/specificity, and OA) were calculated with the testing dataset. OA is the total number of correctly classified segments divided by the total number of test segments. Precision is the true positive rate or producer's accuracy (PA), and recall is the true negative rate or user's accuracy (UA). Depending on whether the dataset is balanced or imbalanced, different performance measures and estimation methods will show different performance when they are applied to supervised algorithms. To select the best classifier, this study compared the performance of the selected four classifiers based on the classification results.

In machine learning, a confusion matrix is a standard means of measuring the performance of a classification model in depth. Most classifiers produce multiple categorical outputs. However, the majority of error measures calculate the total error in the model, not the individual instances of errors. This poses a problem as the model might misclassify some categories more than others, which would not be identified by a standard accuracy measure. Therefore, a confusion matrix is particularly useful for imbalanced data, such as those in this study. The OA, precision, recall, F1 score, and K coefficient were calculated from the confusion matrix to compare the classified maps' accuracy and the performance of the four classifiers. To select the best classifier for the study, the classifiers were also compared using F1 score because the field training samples were imbalanced. The F1 score is a summary measure of the precision and recall, which is calculated as the weighted harmonic mean of both metrics (Ceri et al., 2013):

$$\mathrm{F1} = \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{2.2}$$

Metrics obtained from the confusion matrix were employed for the evaluation of the accuracy of the classification algorithms. To construct the confusion matrix, 30% ground truth samples were used. These

samples were randomly selected before definition of the classifiers; this means that these samples were used to fit the classification algorithms. OA, K coefficient, precision, recall, and F1 score were the metrics used for classification accuracy assessment. The OA, which is for general evaluation, is computed through the sum of correctly classified pixels divided by the validation set's total pixels. The precision, recall, and F1 score metrics were employed to determine the classification quality for each LC class. The PA (precision) indicates that pixels for a class have been omitted, the UA (recall) represents the commission of pixels, and F1 score is the weighted harmonic mean of precision and recall, and thus it represents the degree of accuracy with which the model classifies each class. Like the OA, the K coefficient is a general measure of evaluation, but it is a more representative metric since it also considers wrongly classified samples.

2.2.4.9 Time-effort Investment

Immense time and effort are invested in obtaining acceptable classification accuracies. However, most studies rarely report on the preparation time and the effort required to set up the classification process (e.g., the preparation of the "signature/reference" file), with most work focusing on the time effort for algorithm classification processing time (Nitze, I. et al., 2012; Laborte et al., 2010).

A great deal of effort has also been put into to evaluating the performance of LC classification algorithms (Huang et al., 2002; Pal and Mather, 2005; Rodríguez Galiano et al., 2012). However, few studies have specifically compared algorithms with time-series data (Jia et al., 2014b; Schneider, 2012) and included the amount of effort required in obtaining that classification.

Finally, most researchers focus on the classification accuracy of an algorithm in terms of the confusion matrix values (OA, PA, UA, F1 score, and K coefficient), ignoring the time and effort spent in achieving these accuracies. Therefore, in order to obtain a more complete picture, for this research, metrics of time and effort used in the classification process were also evaluated as part of determining the accuracy of the results.

2.3 Results and Discussion

2.3.1 Visual Examination of Classification Results

The LC classification maps for the Landsat 8 image from September 2018 using the four algorithms are presented in Figure 2.4. The results of two machine-learning non-parametric classifiers (SVM, ANN) are shown in Figures 2.4a and 2.4d, and Figures 2.4b and 2.4c show the classification results of the two traditional parametric classifiers (MLC, MahD). Visual inspection of these classified maps confirms that,

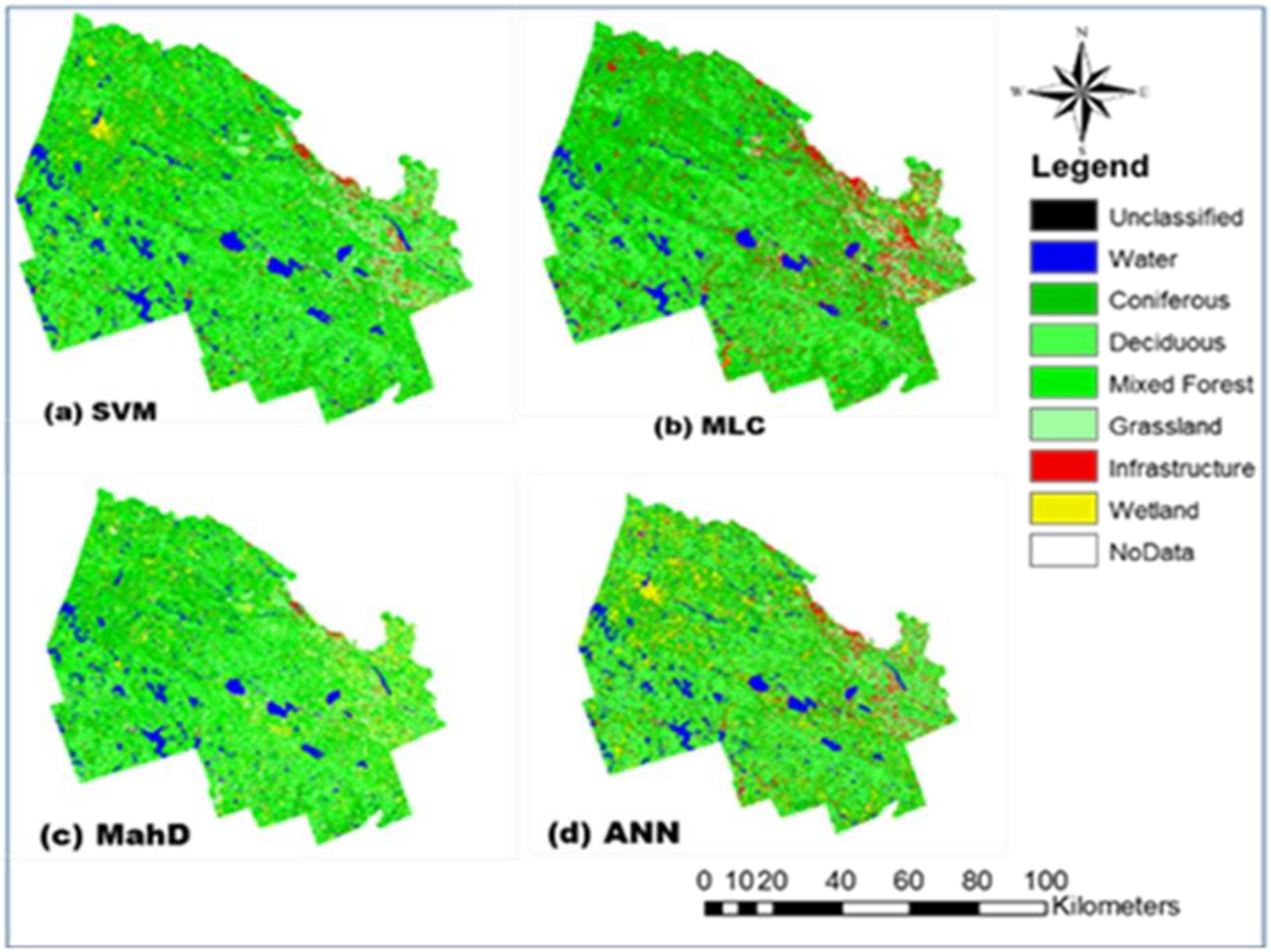


Figure 2. 4 (a-d): LC classification maps created by (a) SVM, (b) MLC, (c) MahD, and (d) ANN classification algorithms

in general, all classifiers can produce functional LC maps and to a large extent, reliable classification results. In terms of these maps, however, it is important to remember they have been affected by limitations related to field sampling, including limited resources and personnel, time, expense, extent of

the study area, and problems of access related to terrain, weather, and risk of encountering wild animals. This means that the ability to extrapolate these accuracy values to the whole forested study area in Ontario should be examined carefully.

2.3.2 Statistical Accuracy Assessment

Confusion matrices resulting from the application of the four algorithms are presented in Tables 2.3 to 2.6. In each case, the same seven land cover classes are considered. For the MahD classifier, an eighth "No Data" class was also needed. Moreover, the ANN classifier was unable to produce results for the Mixed Forest class; instead, it split samples from this class between the Coniferous and Deciduous classes. In terms of accuracy reflected by the confusion matrix, to the right of each matrix are three columns of additional statistics for each class. "Precision" is the fraction of correct predictions made by the classifier, transferred directly from the table. "Recall" represents the sensitivity of the model's ability to avoid false negatives. Finally, the F1 score, which uses a weighted harmonic mean of the Precision and Recall values, gives the most complete overall assessment of classifier accuracy. At the bottom of the tables are additional aggregate values for OA and the K coefficient. Tables 2.3 to 2.6 below summarize the results of the four algorithms.

Table 2. 3: Confusion Matrix (including Precision, Recall, OA, K Coefficient, and F1 Score) of the SVM Classifier (V8)

Classes	Water	Conifer	Deciduous	Mixed Forest	Grassland	Infrastructure	Wetland	Total	Precision	Recall	F1 Score
Water	**100**							100	1.000	1.000	1.000
Conifer		**99.08**					4.07	103.2	0.991	0.961	0.975
Deciduous			**97.5**					97.5	0.975	1.000	0.987
Mixed Forest		0.46	2.5	**100**				103	1.000	0.971	0.985
Grassland					**100**			100	1.000	1.000	1.000
Infrastructure						**97.83**		97.83	0.978	1.000	0.989
Wetland		0.46				2.17	**95.93**	98.56	0.959	0.973	0.966
Total	100	100	100	100	100	100	100	700	6.903	6.905	6.903
Average									**0.986**	**0.986**	**0.986**
OA		**98.62%**									
K Coefficient		**0.984**									

Table 2. 4: Confusion Matrix (including Precision, Recall, OA, K Coefficient, and F1 Score) of the ML Classifier

Classes	Water	Conifer	Deciduous	Mixed Forest	Grassland	Infrastructure	Wetland	Total	Precision	Recall	F1 Score
Water	**87.83**							87.83	0.878	1.000	0.935
Conifer		**92.17**		12.94			17.65	122.8	0.922	0.751	0.828
Deciduous			**87.5**	3.53				91.03	0.875	0.961	0.916
Mixed Forest		5.07	10	**80**				95.07	0.800	0.841	0.820
Grassland			1.25		**81.43**			82.68	0.815	0.985	0.892
Infrastructure		2.3	1.25	3.53	18.57	**100**	11.31	137	1.000	0.730	0.844
Wetland	12.17	0.46					**70.14**	82.77	0.708	0.847	0.771
Total	100	100	100	100	100	100	100	700	5.997	6.116	6.006
Average									**0.857**	**0.874**	**0.865**
OA		**85.69%**									
K Coefficient		**0.833**									

Table 2. 5: Confusion Matrix (including Precision, Recall, OA, K Coefficient, and F1 Score) of the MahD Classifier

Classes	Water	Conifer	Deciduous	Mixed Forest	Grassland	Infrastructure	Wetland	NoData	Total	Precision	Recall	F1 Score
Water	**95.1**					2.9	1.36		99.3	0.965	0.957	0.961
Conifer	0.29	**80.65**		8.24			10.86		100	0.807	0.806	0.806
Deciduous		0.46	**75**	17.65	1.43	0.72			95.3	0.759	0.787	0.773
Mixed Forest		11.52	21.25	**68.24**	1.43				102	0.773	0.666	0.716
Grassland		0.46	1.25		**87.14**	0.72	1.36		90.9	0.871	0.958	0.913
Infrastructure						**83.33**			83.3	0.878	1.000	0.935
Wetland		6.91	2.5	4.71	10	6.52	**79.64**		110	0.775	0.722	0.748
No Data	4.64			1.18		5.8	6.79	**100**	118	1.000	0.845	0.916
Total	100	100	100	100	100	99.99	100.01	100	800	6.829	6.742	6.767
Average										**0.854**	**0.843**	**0.848**
OA		**83.63%**										
K Coefficient		**0.813**										

Table 2. 6: Confusion Matrix (including Precision, Recall, OA, K Coefficient, and F1 Score) of the ANN Classifier

Classes	Water	Conifer	Deciduous	Mixed Forest	Grassland	Infrastructure	Wetland	Total	Precision	Recall	F1 Score
Water	**100**						1.81	101.81	1.000	0.982	0.991
Conifer		**70.97**		43.53				114.5	0.710	0.620	0.662
Deciduous			**100**	52.94				152.94	1.000	0.654	0.791
Mixed Forest				**0**				0	0.000	0.000	0.000
Grassland					**68.57**			68.57	0.686	1.000	0.814
Infrastructure					31.43	**100**	0.45	131.88	1.000	0.758	0.863
Wetland		29.03		3.53			**97.74**	130.3	0.977	0.750	0.849
Total	100	100	100	100	100	100	100	700	5.373	4.764	4.968
Average									**0.768**	**0.681**	**0.721**
OA		**76.75%**									
K Coefficient		**0.729**									

Figures 2.5(a) and 2.5(b) present the various aggregate scores graphically, while Figure 2.6(a–d) does the same broken down by LC class.

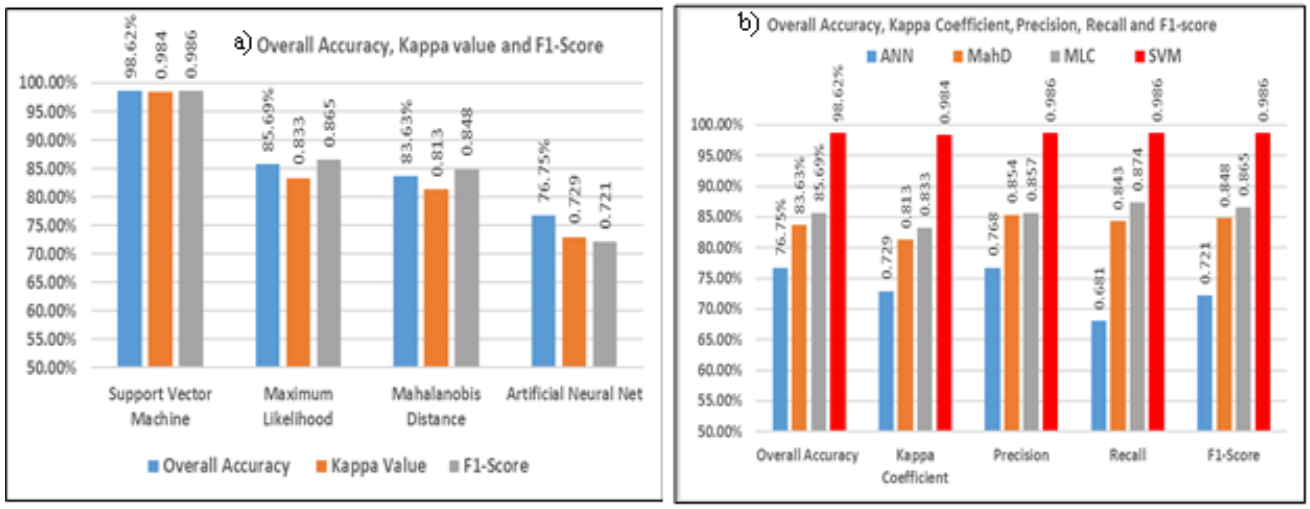


Figure 2. 6 (a-b): (a) OA, K coefficient, and F1 score of the four algorithms compared; (b) OA, K coefficient, average precision, average recall, and average F1 score for the four algorithms compared

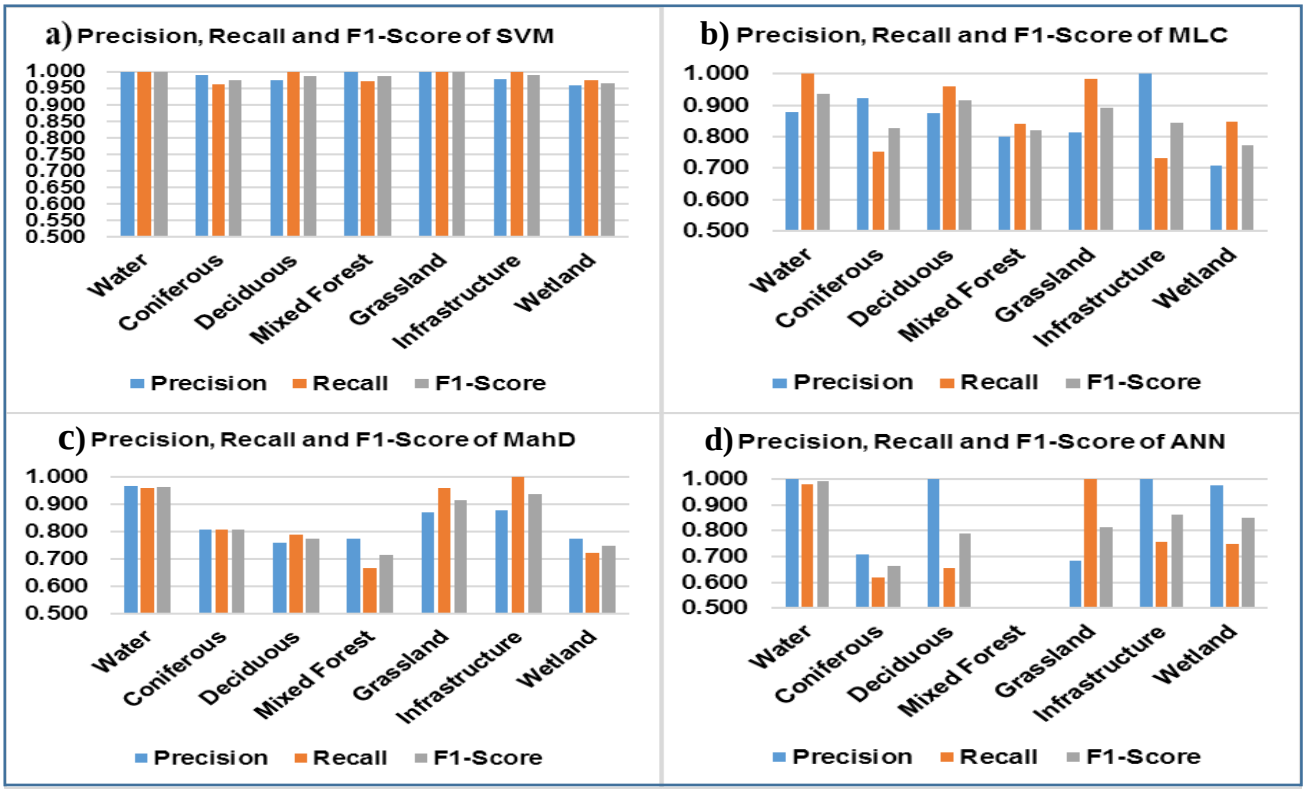


Figure 2. 5 (a–d): Precision, recall, and F1 score of individual classes of the algorithms

2.3.3 Time-effort Investment

An important contribution of this study is the examination of the time and effort invested in obtaining the best classification possible by each of the algorithms compared, since increasing the number of versions of the classification effort would continue to improve the classification accuracy, but it may be practically inoperable due to excess of effort and time invested. Table 2.7 shows the time effort invested for training the signature files used in producing the different versions of classification results.

Table 2. 7: Time-effort investment and OA of different versions

Versions	Time Invested (Hrs)	Overall Accuracy
V1	14.00	60.86%
V2	20.00	72.74%
V3	25.25	85.65%
V4	29.00	90.50%
V5	32.25	95.13%
V6	35.25	96.06%
V7	38.75	97.42%
V8	42.00	98.62%

To increase the classification accuracy, eight versions (V1–V8) of signature files were created to train the cloud-free Landsat image from September 2018. Signature file V1 was created using 175 ground truth samples with 14 LC class types and a color code/scheme was assigned for each signature. It took 14 hours to create the signature file (Table 2.7). The Landsat image of the study area was trained with the V1 signature file using the best classifier, SVM. The OA was 60.86% (see Table 2.8), with some classes poorly classified, even asphalt and dead tree classes were misclassified (false positive). The classification results of the 14 classes were combined into 7 classes ("NoData" class removed from the tables). Buildings, asphalt, and sandy soil were included in the building class and renamed infrastructure; swamp, marsh, bog, and dead tree were combined into swamp and renamed wetland; and coniferous and shelterwood were made into one coniferous class.

Table 2. 8 Confusion Matrix (including Precision, Recall, OA, K Coefficient, and F1 Score) of version 1 (V1)

Class	Water	Coniferous	Deciduous	Mixed Forest	Grassland	Infrastructure	Wetland	Total	Precision	Recall	F1-Score
Water	**99.71**							99.7	0.997	1.000	0.999
Coniferous		**52.07**		2.35			0.45	54.9	0.521	0.949	0.672
Deciduous			**23.75**					23.8	0.238	1.000	0.384
Mixed Forest				**2.35**				2.35	0.024	1.000	0.046
Grassland					**48.57**			48.6	0.486	1.000	0.654
Infrastructure				1.18	8.57	**100**		110	1.000	0.911	0.954
Wetland	0.29	47.93	76.25	94.12	42.86		**99.55**	361	0.996	0.276	0.432
Total	100	100	100	100	100	100	100	700	4.260	6.136	4.140
Average									**0.609**	**0.877**	**0.718**
Overall Accuracy		**60.86%**									
Kappa Coefficient		**0.543**									

To address the misclassified LC class types, the field was revisited before editing the signature files; however, the new samples were not included in the total of 250 samples. The V1 signature file was edited with 13 classes after deleting the dead tree signature class, which required 6 additional hours of effort (without time/hrs for revisiting the field), and it was renamed V2. The classification results of the 13 classes were recombined into 7 classes. Buildings, asphalt, and sandy soil were combined into the building class and renamed infrastructure; swamp, marsh, and bog were classed as swamp and renamed wetland; and coniferous and shelterwood were combined into a single coniferous class. The classification result of V2 was 72.74% (OA), with an increased precision for all forest classes and some other classes, except asphalt, which was misclassified as a building class (false positive).

The field was also revisited for V3 and V4 (see Table 2.9 below), which were edited with 10 classes by merging building and asphalt samples with infrastructure signature and shelterwood samples with coniferous signature. There were 18 and 23 coniferous samples in the V3 and V4 signature files, respectively. Both 5.25 hrs and 3.75 hrs of additional invested time effort were required for the V3 and V4 signature files, respectively (not including time revisiting the field). Infrastructure and sandy soil were combined into the infrastructure class; swamp and marsh were included in swamp and renamed wetland. The OA of V3 and V4 was 85.65% and 90.50%, respectively, with increased accuracies of the individual classes.

Table 2. 9: Confusion Matrix (including Precision, Recall, OA, K Coefficient, and F1 Score) of version 4 (V4)

Class	Water	Coniferous	Deciduous	Mixed Forest	Grassland	Infrastructure	Wetland	Total	Precision	Recall	F1-Score
Water	100							100	1.000	1.000	1.000
Coniferous		98.16		10.59			5.43	114.2	0.982	0.860	0.917
Deciduous			96.25		2.86			99.11	0.963	0.971	0.967
Mixed Forest		0.46	3.75	83.53				87.74	0.835	0.952	0.890
Grassland					61.43			61.43	0.722	1.000	0.838
Infrastructure		0.92		5.88	35.71	100	0.45	143	1.000	0.588	0.741
Wetland		0.46					94.12	94.58	0.941	0.995	0.967
Total	100	100	100	100	100	100	100	700	6.442	6.366	6.320
Average									**0.920**	**0.909**	**0.915**
Overall Accuracy		**90.50%**									
Kappa Coefficient		**0.889**									

The V5, V6, V7, and V8 signature files were edited with 9 signature classes by merging sandy soil samples into the infrastructure signature file after again revisiting the field to collect further samples. The additional effort in terms of time invested for V5, V6, V7, and V8 was 3.25, 3.0, 3.5, and 3.25 hours, respectively, and the cumulative total time effort for V1 to V8 was 14, 20, 25.25, 29, 32.25, 35.25, 38.75, and 42 hours, respectively (Table 2.7). The swamp and marsh classes were combined into the swamp class and renamed wetland. The OA of V5, V6, V7, and V8 was 95.13%, 96.06%, 97.42%, and 98.62%, respectively (see Table 2.7). The increased accuracy from V1 (60.86%) to V8 (98.62%) was 37.76% (Tables 2.3, 2.8, and 2.10). V8 had an average precision, recall, and F1 score of 0.986 (Table 2.3). The classification results obtained from the Landsat images for 1972–2020 with the best algorithm, SVM, used the version 8 signature file (see Table 2.10 below).

Table 2. 10: Details of Landsat imagery including OA of SVM with version 8 reference/signature file and explained variability by the first principal component (PC1)

Acquisition Year	Landsat Platform	Landsat Sensors	Overall Accuracy-V8	Explained Variability (PC1)
2020	LC8	OLI_TIRS	90.80%	94.35%
2019	LC8	OLI_TIRS	89.21%	93.34%
2018	LC8	OLI_TIRS	98.62%	93.02%
2017	LE7	ETM+	88.16%	92.99%
2016	LC8	OLI_TIRS	92.33%	93.71%
2015	LC8	OLI_TIRS	67.97%	92.74%
2014	LC8	OLI_TIRS	91.39%	93.08%
2013	LC8	OLI_TIRS	91.86%	91.72%
2012	LE7	ETM+	94.55%	91.22%
2011	LT5	TM	90.14%	89.17%
2010	LE7	ETM+	89.01%	90.19%
2009	LT5	TM	93.16%	88.35%
2007	LE7	ETM+	90.47%	93.75%
2006	LE7	ETM+	92.03%	92.66%
2005	LT5	TM	94.18%	95.84%
2004	LE7	ETM+	91.66%	96.43%
2003	LE7	ETM+	86.90%	91.67%
2002	LT5	TM	88.06%	96.10%
2001	LT5	TM	95.58%	96.00%
1998	LT5	TM	90.30%	99.35%
1996	LT5	TM	84.60%	99.31%
1995	LT5	TM	93.69%	99.36%
1994	LT5	TM	93.99%	99.43%
1993	LT5	TM	93.64%	99.46%
1992	LT5	TM	91.17%	99.44%
1991	LT5	TM	89.32%	99.33%
1990	LT5	TM	92.24%	99.33%
1987	LT5	TM	93.39%	99.32%
1986	LT5	TM	94.34%	99.32%
1985	LT5	TM	87.39%	99.40%
1984	LT5	TM	92.03%	99.25%
1981	LM2	MSS	94.26%	99.75%
1980	LM2	MSS	92.84%	99.75%
1979	LM2	MSS	82.03%	99.71%
1978	LM3	MSS	89.42%	99.73%
1976	LM2	MSS	84.56%	99.64%
1975	LM1	MSS	86.48%	98.73%
1973	LM1	MSS	85.75%	99.18%
1972	LM1	MSS	94.99%	99.72%

Note: Some years are missing since there is no available and/or usable Landsat image data for those dates. Please note that the explained variability (PC1), which was calculated using principal component analysis, was conducted for a separate study (PhD dissertation), meaning the details are not included this paper (details in chapter 3).

The graphical representation of the classification accuracy with invested time-effort and time-effort for each version are presented in Figure 2.7(a) and 2.7(b), respectively. The classification results in this research show that to obtain the highest acceptable accuracy requires the investment of more time, effort, and cost.

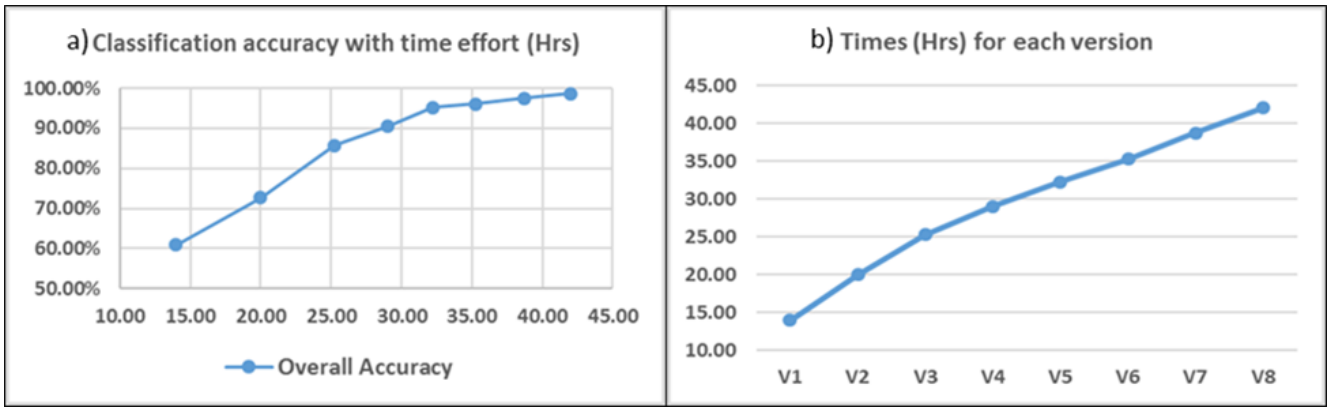


Figure 2. 7 (a-b): a) Classification accuracy with invested time-effort, b) Invested time-effort for each version

In this regard, the increasing degree of accuracy with signature/reference file improvements meant that the time and effort were worthwhile. In other words, the continuous improvements in the "signature/reference file," which were made for training the algorithm, led to increased gains in terms of accuracy. However, it was unclear whether these gains were always proportional or led to diminishing returns in accuracy given the amount of effort.

2.4 Discussion

The results of the statistical analysis in this study showed that SVM is the most accurate algorithm for Landsat image classification, with an OA of 98.62%, a K coefficient of 0.984, and an average F1 score of 0.986. The SVM algorithm outperformed the other three algorithms for all the metrics tested here. Also, SVM was the only algorithm capable of classifying both water and grassland with 100% (F1 score=1.0) accuracy. It should be noted that the SVM algorithm is able to use fewer training samples to achieve a higher level of accuracy compared to other algorithms (Mountrakis et al., 2011).

All LC classes were easily separable by all implemented classifiers, and all classes showed a high level of accuracy (>0.5 score), except for the ANN classifier, which could not read the mixed forest class (0%). Overall, SVM achieved the best classification results when compared with those of the other approaches. MLC displayed the next-best performance, followed by MahD, and ANN had the worst performance. Each algorithm showed distinct behavior for the various classes.

For every classification run, the algorithm was trained (and evaluated) with the training and testing datasets. Training data are an important element of supervised learning since they may affect the expected results negatively or positively based on the training data characteristics and the adaptation of the learning algorithm to the data. Imbalanced training samples are often the result of difficulties in obtaining ground truth data. For this research, the impact of the imbalanced training dataset on the accuracy of the four different supervised classification algorithms was noticeable.

For LU mapping purposes, one of the advantages of the SVM algorithm is that it produces highly accurate classified images from relatively small training datasets (Mountrakis et al., 2011; Salberg and Jenssen, 2012). The experimental results suggested that MLC, MahD, and ANN classifications were affected by imbalanced training data, resulting in a lower level of accuracy, while SVM was not significantly affected by the training data. Thus, SVM was proven to be a very robust, efficient, and consistent classifier since it has the capacity to perform very well with imbalanced training data.

Additionally, SVM exhibited high computational speed, compared to the other algorithms, and had appropriate values for optimization. This, in combination with the accuracy results, suggested that SVM is generally the optimal algorithm, followed by MLC, MahD, and ANN in order of accuracy. However, caution should be taken, as performance depends on a broad search of the parameter space. Additionally, the selection of the training sample composition has a profound impact on the accuracy obtained, meaning it is necessary for users to consider accuracy priorities in their sampling design.

One of the important objectives of this paper was to increase the classification accuracy beyond 90%. To this end, the field was revisited continuously to collect more ground truth data to review potentially misclassified classification factor (LC types) areas until reaching the highest accuracy level. However, information about the number of times the field was revisited and the hours spent there, as well as the number of re-collected samples (which were not included in the total of 250 samples), were not included in the time and effort invested for the signature/reference files. For SVM, the amount of time and effort invested led to gains, as the classification accuracy was increased from 60.86% in V1 to 98.62% in V8, with a time investment of 14hrs and 42hrs, respectively.

The findings of this study are important as they show that SVM has significant advantages for LC classification with Landsat data for a forested area over other classification algorithms. Its high computational speed, the fact that it is not affected by imbalanced training data, and its production of highly accurate classified images from relatively small training datasets makes it a good classification algorithm choice, when available. Moreover, SVM will be useful in future studies in key areas examining the effects of climate change on forest disturbance and degradation. Ontario forests are already being affected by natural and anthropogenic disturbances causing forest degradation, which negatively impacts local ecosystem health, safety, and economy. Therefore, this study has shown that SVM could play a key role in developing a robust methodology using existing RS data and technology for the continuous assessment, mapping, and monitoring of present and historic forest disturbance and degradation trends as well as their projections under climate change.

2.5. Conclusions

In this study, the classification of Landsat 8 imagery was undertaken with two different advanced non-parametric machine learning algorithms—SVM and ANN—as well as the two parametric traditional classification algorithms—MLC and MahD. These algorithms were implemented, evaluated, and

compared. The results were assessed in terms of LC classification accuracy, based on the use of Landsat image data from the study area for classification. Additionally, for classification assessment, the ground truth points approach was employed to determine the OA, K coefficient, and F1 score.

SVM with an OA of 98.62%, K coefficient of 0.984, and F1 score of 0.986 is more accurate than the other algorithms. Additionally, SVM's accuracy was not affected by imbalanced training data. In terms of individual class accuracies, SVM showed the highest accuracy for all classes (>95%). The findings revealed that SVM is a very robust, consistent, and effective classifier as it performs very well with an imbalanced dataset. While a greater amount of time and effort was spent (hours) in making improvements to the reference file to increase the accuracy from 60.86% (14 hrs) to 98.62% (42 hrs) with the SVM algorithm, it was unclear whether the gains were truly proportional to the extra time, effort, and expense. The results presented in this paper can be employed as the foundation for future research into satellite image classification of LC. First, future work would benefit from using more field data, whose collection was limited in this study given issues of accessibility. The collection of ground truth/field data proved to be expensive, time-consuming, and difficult, given the large study area, as well as the time spent preparing the reference/signature files. Subsequent studies should also center on algorithms with high accuracy to achieve the optimal parameters to ensure the accurate classification of Landsat satellite images. However, classifiers with balanced training field data, as well as fine resolution image data for training, should be investigated further when using Landsat imagery. The use of SVM on this dataset can improve the results owing to its capacity for the engineering of new features. Finally, further work should also compare the accuracy and computational complexities of the advanced machine learning algorithms in terms of the time and effort necessary to achieve such accuracies.

Disclosure statement

The authors declared no conflicts of interest.

References

Aguilar, M.A., M.M. Saldaña, F.J. Aguilar. 2012. “GeoEye-1 and WorldView-2 pan-sharpened imagery for object-based classification in urban environments.” *International Journal of Remote Sensing*, 34; 2583–2606.

Aguirre-Salado, C., E. Trevino-Garza, O. Aguirre-Calderon, J. Jimenez-Perez, M. Gonzalez-Tagle, L. Miranda-Aragon, J. Valdez-Lazalde, A. Aguirre-Salado, and G. Sanchez-Diaz. 2012. “Forest Cover Mapping in North-Central Mexico: A Comparison of Digital Image Processing Methods.” *GIScience and Remote Sensing* 49 (6): 895–914.

Bailly, J.S., Arnaud, M., Puech, C. 2007. “Boosting: a classification method for remote sensing.” *Int. J. Remote Sens.* 28, 1687–1710.

Bakr, N., D.C. Weindorf, M.H. Bahnassy, S.M. Marei, and M.M. El-Badawi. 2010. “Monitoring land cover changes in a newly reclaimed area of Egypt using multi-temporal Landsat data.” *Applied Geography* 30:592-605.

Ballanti, L.; Blesius, L.; Hines, E.; Kruse, B. 2016. “Tree species classification using hyperspectral imagery: A comparison of two classifiers.” *Remote Sens.* 8, 445.

Bhattarai, N., Quackenbush, L. J., Dougherty, M., & Marzen, L. J. 2015. “A simple Landsat–MODIS fusion approach for monitoring seasonal evapotranspiration at 30 m spatial resolution.” *International Journal of Remote Sensing*, 36(1), 115–143.

Brereton, R.G. and Lloyd, G.R. 2010. “Support vector machines for classification and regression.” *Analyst*. 135, 230–267.

Brown, M., Gunn, S.R., Lewis, H.G. 1999. Support vector machines for optimal classification and spectral unmixing. *Ecol. Model.*, *120*, 167–179.

Buckland, M. and Gey, F. 1994. “The Relationship Between Recall and Precision.” *Journal of the American Society for Information Science*, 45(1): 12-19.

Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Rojo-Alvarez, J.L. and Martinez-Ramon, M. 2008. "Kernel-Based Framework for Multitemporal and Multisource Remote Sensing Data Classification and Change Detection," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 6, pp. 1822-1835.

Ceri, S., Bozzon, A., Brambilla, M., Della Valle, E., Fraternali, P., Quarteroni, S. 2013. *An Introduction to Information Retrieval.* Springer: Berlin/Hiedelberg, Germany; pp. 3–11.

Chaichoke, V, P. Supawee, V. Tanasak and K.S. Andrew. 2011. “A Normalized Difference Vegetation Index (NDVI) time-series of idle agriculture lands: A preliminary study.” *Engineering Journa*l, 15(1): 9-16.

Chubey, M., Stehle, K., Albricht, R., Gougeon, F., Leckie, D., Gray, S., Woods, M., Courville, P. 2009. “Semi-automated species classification in Ontario Great Lakes – St. Lawrence forest conditions. Final Report: Great Lakes - St. Lawrence ITC Project (2005/2008).” *Ontario Ministry of Natural Resources*, January 2009, p71.

Cihlar, J. 2000. “Land cover mapping of large areas from satellites: status and research priorities.” *International Journal of Remote Sensing*, 21 (6-7), 1093–1114.

Congalton, R.G. and Green, K. 2019. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, 3rd Ed.; CRC Press: Boca Raton, FL, USA.

Dixon, B. and Candade, N. 2008. “Multispectral land use classification using neural networks and support vector machines: one or the other, or both?” *International Journal of Remote Sensing*, 29; 1185–1206.

Eliason, E.M. and McEwen, A.S. 1990. “Adaptive Box Filters for Removal of Random Noise from Digital Images.” *Photogrammetric Engineering & Remote Sensing*, 56 (4): 453-458.

Environment for Visualizing Images (ENVI) version 5.3. Exelis Visual Information Solutions, Boulder, CO, USA. http://www.harrisgeospatial.com/ProductsandSolutions/GeospatialProducts/ENVI.

Favot, E.J., Rühland, K.M., DeSellas, A.M. *et al.* 2019. Climate variability promotes unprecedented cyanobacterial blooms in a remote, oligotrophic Ontario lake: evidence from paleolimnology. *J Paleolimnol* 62, 31–52.

Foody, G.M. 2002. “Status of land cover classification accuracy assessment.” *Remote Sens. Enviro*, 80, 185–201.

Foody, G.M. and Mathur, A. 2004. “A relative evaluation of multiclass image classification by support vector machines.” *IEEE Transactions on Geoscience and Remote Sensing*, 42:1336–1343.

Foody, G.M., Mathur, A., Sanchez-Hernandez, C., Boyd, D.S. 2006. “Training set size requirements for the classification of a specific class.” *Remote Sensing of Environment*, 104, 1-14.

Foody, G.M., McCulloch, M.B. and Yates, W.B. 1995. “The effect of training set size and composition on artificial neural network classification.” *International Journal of Remote Sensing*, 16, 1707–1723.

Ghayour, L.; Neshat, A.; Paryani, S.; Shahabi, H.; Shirzadi, A.; Chen, W.; Al-Ansari, N.; Geertsema, M.; Pourmehdi Amiri, M.; Gholamnia, M.; et al. 2021. “Performance Evaluation of Sentinel-2 and Landsat 8 OLI Data for Land Cover/Use Classification Using a Comparison between Machine Learning Algorithms.” *Remote Sens*. 13, 1349.

Ghosh, A.; Joshi, P.K. 2014. “A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery.” *Int. J. Appl. Earth Obs*. Geoinf. 26, 298–311.

Hall, F.G., Strebel, D.E., Nickeson, J.E and Goetz, S.J. 1991. “Radiometric rectification: Toward a common radiometric response among multidate, multisensor images,” Remote Sens. Environ., vol. 35, pp. 11–27, 1991.

Hansen, M. C. and Loveland, T. R. 2012. “A review of large area monitoring of land cover change using Landsat data.” *Remote Sens. Environ*. 122, 66–74.

Hansen, T. 2012. A review of large area monitoring of land cover change using Landsat data. Remote Sens. Environ. 122, 66-74.

Haralick, R. M; Sternberg, S. R; and Zhuang X. 1987. “Image Analysis Using Mathematical Morphology.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-9, No. 4, July 1987, pp. 532-550.

Huang, C., Davis, L.S., Townshend, J.R.G. 2002. “An assessment of support vector machines for land cover classification.” *Int. J. Remote Sens*. 23, 725–749.

Huang, C., Song, K., Kim, S., Townshend, J.R.G, Davis, P., Masek, J.G, Goward, S.N. 2008. “Use of a dark object concept and support vector machines to automate forest cover change analysis.” *Remote Sens Environ.* 112: 970–985.

Jensen, J. 2005. *Introductory digital image processing: A remote sensing perspective (3rd ed.)*. Upper Saddle River, NJ: Prentice Hall.

Jia, K., Liang, S., Zhang, L., Wei, X., Yao, Y. and Xie, X. 2014. Forest cover classification using Landsat ETM+ data and time series MODIS NDVI data. *International Journal of Applied Earth Observation and Geoinformation, 33,* pp.32-38.

Jog, S. and Dixit, M. 2016. "Supervised classification of satellite images." 2016 Conference on Advances in Signal Processing (CASP), pp. 93-98.

Kalantar, B., Pradhan, B., Naghibi, S.A., Motevalli, A., Mansor, S. 2018. "Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN)." *Geom. Nat. Haz. Risk* 9 (1), 49–69.

Kar, S.A. and Kelkar, V.V. 2013. Classification of Multispectral Satellite Images, International Conference on Advances in Technology and Engineering (ICATE).

Kavzoglu, T. 2009. "Increasing the accuracy of neural network classification using refined training data." *Environmental Modelling & Software*, 24, 850-858.

Kavzoglu, T. and Mather, P.M. 2003. "The use of backpropagating artificial neural networks in land cover classification." *International Journal of Remote Sensing*, 24, pp. 4907–4938.

Kavzoglu, T., and Colkesen, I. 2009. "A Kernel Functions Analysis for Support Vector Machines for Land Cover Classification." International Journal of Applied Earth Observation and Geoinformation 11: 352–359.

Khatami, R., Mountrakis, G., Stehman, S.V. 2016. "A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: general guidelines for practitioners and future research." *Remote Sens. Environ*. 177, 89–100.

Knorn, J., Rabe, A., Radeloff, V.C., Kuemmerle, T., Kozak, J., Hostert, P. 2009. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote. Sens. Environ.*, *113*, 957–964.

Köppen, Wladimir. 1936. "C". In Köppen, Wladimir; Geiger (publisher), Rudolf (eds.). Das geographische System der Klimate [The geographic system of climates] *(PDF)*. Handbuch der Klimatologie. 1. Berlin: Borntraeger. Archived *(PDF)* from the original on 2016-03-04. *Retrieved 2016-09-02*.

Laborte, A.G., Maunahan, A.A., Hijmans, R.J. 2010. Spectral Signature Generalization and Expansion Can Improve the Accuracy of Satellite Image Classification. PLoS ONE 5(5).

Li, C., Wang, J., Wang, L., Hu, L., Gong, P. 2014. Comparison of classification algorithms and training sample sizes in urban land classification with Landsat Thematic Mapper imagery. *Remote Sens.*, *6*, 964–983.

Lu, D. and Weng, Q. 2007. "A survey of image classification methods and techniques for improving classification performance." *International Journal of Remote Sensing*, 28(5), 823– 870.

Mas, J. F., and J. J. Flores. 2008. "The Application of Artificial Neural Networks to the Analysis of Remotely Sensed Data." *International Journal of Remote Sensing*, 29: 617–663.

McKenney, D.W., Pedlar, J.H., Lawrence, K., Gray, P.A., Colombo, S.J., Crins, W.J. 2010. Current and projected future climatic conditions for ecoregions and selected natural heritage areas in Ontario. Ontario Ministry of Natural Resources, Applied Research and Development Branch, Sault Ste. Marie, ON. Climate change research report CCRR-16

Melgani, F., and Bruzzone, L. 2004. "Classification of hyperspectral remote sensing images with support vector machines." *IEEE Transactions on Geoscience and Remote Sensing*, 42, 1778-1790.

Mohd Pozi, M.S., Sulaiman, M.N., Mustapha, N., & Perumal, T. 2015. “A new classification model for a class imbalanced data set using genetic programming and support vector machines: case study for wilt disease classification.” *Remote Sensing Letters*, 6, 568-577.

Mountrakis, G., Im, J., Ogole, C. 2011. “Support vector machines in remote sensing: A review.” *ISPRS Journal of Photogrammetry and Remote Sensing*, 66; 247–259.

Munyati, C. 2000. “Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset." *International Journal of Remote Sensing*. 21(9) :1787-1806.

Muñoz-Marí, J; Bruzzone, L. and Camps-Valls, G. 2007. "A Support Vector Domain Description Approach to Supervised Classification of Remote Sensing Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 8, pp. 2683-2692.

Nitze, I., Schulthess, U., Asche, H. 2012. Comparison of Machine Learning Algorithms Random Forest, Artificial Neural Network, and Support Vector Machine to Maximum Likelihood for Supervised Crop Type Classification; Proceedings of the 4th GEOBIA, May 7-9, 2012, Rio de Janeiro, Brazil, p.035.

Otukei, J. R., and Blaschke, T. 2010. “Land Cover Change Assessment Using Decision Trees, Support Vector Machines and Maximum Likelihood Classification Algorithms.” International Journal of Applied Earth Observation and Geoinformation 12: S27–S31.

Pal, M., A. E. Maxwell, and T. A. Warner. 2013. “Kernel-Based Extreme Learning Machine for Remote-Sensing Image Classification.” *Remote Sensing Letters*, 4: 853–862.

Pal, M., and Mather, P. M. 2005. “Support Vector Machines for Classification in Remote Sensing.” International Journal of Remote Sensing 26 (5): 1007–1011.

Palaniswami, C., Upadhyay, A. K. and Maheswarappa, H. P. 2006. "Spectral mixture analysis for subpixel classification of coconut", *Current Science*, Vol. 91, No. 12, pp. 1706 -1711.

Pouliot, D., Latifovic, R., Pasher, J., Duffe, J. 2019. “Assessment of Convolution Neural Networks for Wetland Mapping with Landsat in the Central Canadian Boreal Forest Region.” *Remote Sens.* 11, 772.

Powell, R.L., MATZKE, N., DE SOUZA JR, C., CLARK, M., NUMATA, I., HESS, L.L. and ROBERTS, D.A. 2004. “Sources of error in accuracy assessment of thematic land-cover maps in the Brazilian Amazon.” *Remote Sensing of Environment*, 90, pp. 221–234.

Qian, Y.; Zhou, W.; Yan, J.; Li, W.; Han, L. 2015. “Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery.” *Remote Sens.* 7, 153–168.

Qiu, S., He, B., Zhu, Z., Liao, Z., Quan, X. 2017. “Improving Fmask cloud and cloud shadow detection in mountainous area for Landsats 4–8 images.” *Remote Sens. Environ.* 199, 107–119.

Qiu, S., Lin, Y., Shang, R., Zhang, J., Ma, L., Zhu, Z., Qiu, S., Lin, Y., Shang, R., Zhang, J., Ma, L., Zhu, Z. 2019. “Making Landsat time series consistent: evaluating and improving Landsat analysis ready data.” *Remote Sens.* 11 (1), 51.

Qiu, S., Zhu, Z., He, B. 2019. “Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery.” *Remote Sensing of Environment* (in review 2018-published-2019).

Richards J.A. 1999. *Remote Sensing Digital Image Analysis*, Springer-Verlag, Berlin. 240 pp.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., and Rigol-Sanchez, J.P. 2012. “An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification.” ISPRS Journal of Photogrammetry and Remote Sensing 67: 93–104.

Roy, D.P., Wulder, M.A., Loveland, T.R., Carabajal CC, Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., et al. 2014. "Landsat-8: science and product vision for terrestrial global change research." *Remote Sensing of Environment,* 145:154–172.

Salberg, B. and Jenssen, R. 2012. "Land-cover classification of partly missing data using support vector machines." *International Journal of Remote Sensing*, 33; 4471-4481.

Schneider, A. 2012. Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach. Remote Sens. Environ., 124, pp. 689-704.

Sellers, P. J., Meeson, B. W., Hall, F. G., Asrar, G., Murphy, R. E., Schiffer, R. A., Bretherton, F. P., *et al.* 1995. Remote sensing of the land surface for studies of global change: models—algorithms—experiments. *Remote Sensing of Environment*, 51, 3–26.

Shaw, S. B., Marrs, J., Bhattarai, N., & Quackenbush, L. 2014. "Longitudinal study of the impacts of land cover change on hydrologic response in four mesoscale watersheds in New York State, USA." *Journal of Hydrology*, 519, 12–22.

Shetaii S.H, Abdi, O. 2008. "Mapping of land use in mountainous regions of Zagros using ETM+ data." *Journal of Agricultural Sciences and Natural Resources*, University of Gorgan, Gorgan., 57:129-138.

Shi, D., Yang, X. 2015. Support vector machines for land cover mapping from remote sensor imagery. In *Monitoring and Modeling of Global Changes: A Geomatics Perspective*; Springer: Dordrecht, The Netherlands, pp. 265–279.

Shim, D. 2014. "Remote sensing place: Satellite images as visual spatial imaginaries." *Geoforum,* 51; 152–160.

Smeeton, N.C. 1985. "Early history of the kappa statistic." *Biometrics*, 41(3):795-795.

Soffianian, A. and Madanian, M. 2015. "Monitoring land cover changes in Isfahan Province, Iran using Landsat satellite data." *Environmental Monitoring and Assessment,* 187(8), 1–15.

Srivastava, P.K., Han, D., Rico-Ramirez, M.A., Bray, M., T. Islam, T. 2012. "Selection of classification techniques for land use/land covers change investigation." *Advances in Space Research*, 50; 1250–1265.

Stehman, S.V. 1992. "Comparison of systematic and random sampling for estimating the accuracy of maps generated from remotely sensed data." *Photogrammetric Engineering & Remote Sensing*, 58:1343-1350.

Su, L., M. J. Chopping, A. Rango, J. V. Martonchik, and D. P. C. Peters. 2007. "Support Vector Machines for Recognition of Semi-Arid Vegetation Types Using MISR Multi-Angle Imagery." Remote Sensing of Environment 107: 299–311.

Talukdar, S., Singha, P., Mahato, S., Pal, S., Liou, Y.A., Rahman, A. 2020. "Land-use landcover classification by machine learning classifiers for satellite observations—A review." *Remote Sens.* 12 (7), 1135.

Tan, J., Zuo, J., Xie, X., Ding, M., Xu, Z., Zhou, F. 2021. "MLAs land cover mapping performance across varying geomorphology with Landsat OLI-8 and minimum human intervention." *Ecol. Inform.* 61, 101227.

Thanh Noi P. and Kappas M. 2017. "Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery." *Sensors (Basel)*. 18(1):18.

Thessler, S., Sesnie, S, Ramos Bendaña, Z.S, Ruokolainen, K, Tomppo, E. and Finegan, B. 2008. "Using k-nn and Discriminant Analyses to Classify Rain Forest Types in a Landsat TM Image Over Northern Costa Rica." Remote Sensing of Environment 112 (5): 2485–2494.

Tigges, J., Lakes, T., Hostert, P. 2013. "Urban vegetation classification: Benefits of multitemporal RapidEye satellite data." *Remote Sensing of Environment*, 136; 66-75.

USGS. 2016. Product Guide. Provisional Landsat 8 Surface Reflectance Code (LASRC) Product. ''Department of Interior, U.S. Geological Survey" <http://landsat.usgs.gov/CDR_LSR.php>.

USGS. 2019. Landsat 8 (L8) Data Users Handbook Version 5.0. Department of Interior, U.S. Geological Survey.

USGS. 2020. Landsat 4-7 Collection 1 (C1), Surface Reflectance (LEDAPS) Product Guide, Version 3.0. Department of Interior, U.S. Geological Survey.

USGS. 2020. Landsat 8 Collection 1 (C1), Land Surface Reflectance Code (LaSRC) Product Guide, Version 3.0. Department of Interior, U.S. Geological Survey.

Vapnik, V. N. 1998. *Statistical Learning Theory*. New York: Wiley.

Verburg, P.H., Neumann, K., Nol, L. 2011. "Challenges in using land use and land cover data for global change studies." *Glob. Chang. Biol.*, 17, 974–989.

Vermote, E., Justice, C., Claverie, M., & Franch, B. 2016. "Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product." *Remote Sensing of Environment*, 185, 46-56.

Vibhute, A. D; Gawali, B. W. 2013. "Analysis and Modeling of Agricultural Land use using Remote Sensing and Geographic Infor- mation System: a Review", *International Journal of Engineering Research and Applications (IJERA)*, vol. 3(3), pp. 081-091.

Wang, Z. J., Jiao, J. Y., Lei, B., & Su, Y. 2015. "An approach for detecting five typical vegetation types on the Chinese Loess Plateau using Landsat TM data." *Environmental Monitoring and Assessment*, 187(9), 1–16.

Waske, B., Benediktsson, J.A., & Sveinsson, J.R. 2009. "Classifying Remote Sensing Data with Support Vector Machines and Imbalanced Training Data." In: J.A. Benediktsson, J. Kittler & F. Roli (Eds.), *Multiple Classifier Systems: 8th International Workshop, MCS 2009, Reykjavik, Iceland*, June 10-12, 2009. Proceedings (pp. 375-384). Berlin, Heidelberg.

Williams, D.L., Goward, S. and Arvidson, T. 2006. "Landsat: yesterday, today, and tomorrow." *Photogrammetric Engineering and Remote Sensing*, 72:1171–1178.

Wulder, M.A., Franklin, S.E., White, J.C., Linke, J., Magnussen, S. 2006. "An accuracy assessment framework for large-area land cover classification products derived from medium-resolution satellite data." *Int. J. Remote Sens.* 27 (4), 663–683.

Yang, C., Everitt, J.H. and Murden, D. 2011. "Evaluating high resolution SPOT 5 satellite imagery for crop identification." *Computers and Electronics in Agriculture*, 75:347-354.

Yousefi, S., Mirzaee, S., Tazeh, M., Pourghasemi, H., Karimi, H. 2015. "Comparison of different algorithms for land use mapping in dry climate using satellite images: a case study of the Central regions of Iran." *Desert*, 20(1),1-10.

Yu, C., Chen, J. 2020. "Landslide susceptibility mapping using the slope unit for southeastern Helong City, Jilin Province, China: a comparison of ANN and SVM." *Symmetry*, 12 (6), 1047.

CHAPTER THREE

FOREST DEGRADATION INDICATORS AND LONG-TERM TRAJECTORY ANALYSIS

Abstract

During the last few decades, vegetation indices (VIs) calculated from ratioing different wavebands of satellite images have been employed widely for tracking and monitoring different aspects of forest condition, including disturbance and degradation. VIs acquired from remote sensing-based canopies are relatively simple and effective mathematical combinations of image wavebands converted to algorithms for undertaking quantitative and qualitative evaluations of elements such as vegetation/forest cover, vigor, and growth dynamics. They have been broadly used with remote sensing applications that employ various airborne and satellite platforms that have been bolstered by modern advances in unmanned aerial vehicles. In this study, seven VIs were used to derive indicators of forest degradation and attempt to reduce dimensionality by combining them into a composite forest degradation indicator (CFDI). Normalized difference vegetation index (NDVI), net primary productivity (NPP), normalized burn ratio (NBR), enhanced vegetation index (EVI), tasseled cap brightness (TCB), tasseled cap greenness (TCG), and tasseled cap wetness (TCW) from remote sensing (Landsat and MODIS platforms as sources) data were chosen as the indicators that best described forest degradation.

Given the prevalence and availability of image data that are multi-sensor, multi-temporal, multi-resolution, and multifrequency from operational Earth Observation satellites, fusing such digital image data has become an important means of the evaluation of forest condition over time from remote sensing images. Recently, digital satellite image fusion in remote sensing has been able to increase the quality and application of image data. Image fusion involves creating a single image with the relevant information from a set of images. The fused image that results should be more informative and comprehensive than any input image alone. Principal Component Analysis (PCA) is a multivariate statistical analysis

technique used to transform the original image to its eigenspace, reducing dimensionality by means of keeping the principal components (PCs) with influencing eigenvalues (explanatory variance). PCA maintains the crucial features of the original image and decreases the level of unexplained variance or noise.

This study also examined the usefulness of analyzing forest cover change trajectories based on Landsat time series through the application of LandTrendr (Kennedy et al., 2010), a well-known, ad-hoc developed, spectral-temporal segmentation algorithm employed for forest disturbance and recovery detection. This method provides estimates of discontinuous phenomena (in terms of both disturbance date and intensity) and continuous phenomena (such as post-disturbance regeneration) (Kennedy et al., 2010). With the LandTrendr algorithm, alterations in the duration, year, and magnitude of forest disturbances were computed and detected from the aforementioned seven indicators. As part of the following procedure, the inter-annual change results calculated from LandTrendr were combined through a robust estimation procedure based on PCA for reconstructing a change map of annual forest cover in the study area through time-series of imagery.

All FDI variables were combined into a single CFDI for each year, using the image fusion technique reliant on PCA, with PC1 accepted with high explained variability. From 1972–1981, the CFDI variability ranged from 98.73%–99.75%; from 1984–1998, it was from 99.25%–99.46%; and from 2001–2020, it ranged from 88.35%–96.43%. The explained variability of the combined CFDIs of 1972–1981, 1984–1998, and 2001–2020 were 85.01%, 85.69%, and 75.49%, respectively.

Overall, it was found that the innovative CFDI (1972–2020) was the most informative FDI as the combination of three layers of CFDI (1972–1981, 1984–1998, and 2001–2020), and the explained variability of the combined CFDI was 95.32%.

In short, this study developed a fast and effective approach for monitoring forest disturbance and degradation from multiscale and multi-temporal remote sensing data. It demonstrated that multiscale remote sensing data have significant potential when it comes to monitoring forest degradation at the regional level. The results of this thesis showed that Landsat and MODIS are satellite platforms that can be used to derive indicators and a CFDI, which were effective in the identification and mapping of forest disturbance and degradation over the long term.

Keywords: Forest degradation indicators, Landsat time series, principal component analysis, trajectory analysis, LandTrendr algorithm.

3.1 Introduction

Developing efficient methods for monitoring forest degradation is very important both in revealing the process of forest quality reduction, stand structure destruction, and permanent loss of forest and also improving our comprehension of global or regional carbon emissions and the connection with climate change. Recently, advances in time series remote sensing data processing have shown that yearly vegetation indices (VIs) produced by the MODerate Resolution Imaging Spectroradiometer (MODIS) might have a major part in the assessment of forest disturbance and degradation at the regional level (Bradley et al., 2007; Lunetta et al., 2006).

Several studies have attempted to identify forest degradation in terms of both carbon reduction and canopy fraction cover decrease. For analyzing forest health and degradation level, time series of multispectral VIs—including normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI)—which describe vegetation phenology, have been employed in these studies. Other indicators to describe forest productivity, including gross primary production (GPP)/net primary productivity (NPP), have also been adopted to examine the degradation level present in regional areas. However, there were variations in accuracy when employing different indicators and criteria (Romero-Sanchez and Ponce-Hernandez,

2017; Hamunyela, 2015). In addition, the effect of multiscale remote sensing indicators on the effectiveness and sensitivity for detection of forest degradation is as yet uncertain. Examining regional forest degradation from many indicators based on satellite observations data, as well as analyzing the impact of indicators on the accuracy of degraded forest detection, is highly significant in this study.

When it comes to developing a robust method for forest degradation mapping from satellite or airborne remote sensing imagery, challenges remain. In most studies, the effects of climate and its changing trends on forest degradation have not been considered in the detection approaches (Hamunyela, 2015; Lambert et al., 2013). As a result, it is of great importance to develop an approach that includes climate's influence to improve the precision of forest degradation detection using remote sensing. The goal of this study was to establish an indicator system for forest degradation estimation and to analyze the trajectory over a time series of the indicators from the Landsat and MODIS imagery, as well as to develop a method based on multiple indicators as a mean of forest degradation detection in the Ontario study area.

3.1.1 Derived forest degradation indicators (FDIs)

In this study, forest degradation indicators (FDIs) were derived from VIs; in other words, VIs were used as FDIs. VIs are derived from vegetation reflectance properties and constructed from reflectance measurements in two or more wavelengths across the optical spectrum in order to engage in the analysis of certain vegetation characteristics (e.g., water content and total leaf area, etc.). While over 100 VIs derived from band ratioing and other combinations are found in the literature, very few have a significant biophysical basis or have been systematically examined (Asner, 1998 [ENVI^{TM}]). VIs are grouped into categories that calculate similar properties, such as broadband and narrowband greenness indices, light use efficiency indices, and canopy water content indices (Asner, 1998 [ENVI]). However, each category usually supplies multiple techniques for estimation of the absence or presence of one vegetation property. When it comes to different conditions in the field and properties, some indices within a category are able to furnish results with a higher level of validity than others (Asner, 1998 [ENVI]). Estimating vegetation

properties over large areas is efficient with remote sensing, and spectral VIs calculated using ENVI™ software can be employed for analyzing such properties.

VIs involve the development of a numerical relationship between satellite data and a specific biophysical component as the construct of a remote sensing biophysical parameter (Raynolds et al., 2006a). Additionally, VIs are spectral transformations that involve two or more bands that have been designed to enhance vegetation property contributions and allow for the creation of reliable spatial and temporal inter-comparisons of terrestrial photosynthetic activity and variations in tree canopy structure. Given the fact that they involve the simple transformation of spectral bands, VIs are computed directly and must not have any bias or assumptions related to land cover class, type of soil, or climate conditions. A great deal of research has examined the application of Landsat-derived VIs for various types of ecosystems (e.g., Cohen and Goward, 2004; Wulder, 2008b). Owing to their efficiency in estimating many aspects related to plants, such as above-ground biomass, leaf area index (LAI), and primary production, VIs are commonly employed (Chen et al., 2009b; Kushida et al., 2009; Walker et al., 2003). There is proof that VIs enhance vegetation detection (Tucker and Sellers, 1986). Ratio-based VIs are also able to minimize topographic-induced variance (Avery and Berlin, 1992). Generally, VI differencing has outperformed other multi-date methods (e.g., ratioing and image differencing) (Lyon et al., 1998).

Concerning ecological studies, the NDVI is a very widely employed VI. The majority of studies on vegetation change on national to global scales that use satellite remote sensing imagery have only employed the NDVI to serve as a proxy for gross photosynthesis (Chapin, 2005; Vierling et al., 1997). An NDVI value is generally represented as a ratio of surface reflectance from the red and near infrared (NIR) radiometric bands, in which low values denote no vegetation (e.g., water, soil, sand) and high values are vegetation indicators (Campbell, 1996; Lillesand et al., 2008). Gizachew et al. (2016) revealed that in the assessment and monitoring of global carbon balance, forest biomass estimation in complex stand

structures and closed canopy forests could be monitored with accuracy using Landsat-derived NDVI. Employing NDVI as a greenness indicator actually helped to overcome low reflectance values in the NIR that otherwise would have led to saturation problems in the data, thereby rendering it unusable for assessments, and it has been used in Canada for monitoring and assessing boreal forests with evergreen conifers and relatively close canopy forest (Gerylo et al., 2002; Hall et al., 2006).

Time series of multispectral VIs have the potential to describe vegetation phenology changes and help in the monitoring of terrestrial photosynthetic vegetation activity for research related to change detection and biophysical interpretations (Ramon et al., 2010); therefore, NDVI is beneficial since it is able to minimize the band-correlated noise and the influences caused by variations in a variety of elements, including cloud cover and topography. Additionally, EVI has the potential to detect vegetation that is sparsely distributed and decouple the noise caused by aerosol scattering and soil background. It does this through the inclusion of a blue reflectance and feedback term that does the job of simultaneous correction.

When the goal is the detection of low-magnitude disturbances using Landsat time series, the signal associated with spectral change from the disturbance might be masked by noise connected with normal temporal variation resulting from atmospheric and geometric corrections, variations in the angle of the sun, vegetation phenology, and sensor degradation. Since most recent algorithms use a limited number of spectral bands or indices to detect disturbances, such as the "forestness index" (Huang et al., 2010), NBR (Kennedy et al., 2012), and tasseled cap angle (TCA) (Brook et al., 2014), it is crucial to take into account the comparative signal-to-noise strengths among these spectral indices.

While NDVI does have a number of drawbacks, it remains useful for coupling climate and vegetation at large spatial scales (Pettorelli et al., 2005), making it popular in many studies related to land surface dynamics. NDVI has been shown to correlate directly with vegetation productivity (Reed et al., 1994) and been linked with fraction of absorbed photosynthetically active radiation (fAPAR) (Sellers et al., 1992;

Rautiainen et al., 2010). In many ecosystems, NDVI can be employed to determine the direct impacts of climate conditions on vegetation biomass and overall biomass (Nemani et al., 2003; Roerink et al., 2003; Zhou et al., 2003). This would allow for the assessment of the feedback effect vegetation has on the local climate (Jingyong et al., 2003). Additionally, NDVI has been employed to better predict and assess the impact of disturbances (e.g., drought, fire, flood, frost). However, time series of other variables that are relevant (e.g., LAI, fAFAR) and other vegetation indices (e.g., EVI, PPI) could also be beneficial in improving our comprehension of vegetation dynamics.

As an optimized index, the EVI (Huete et al., 2002) was developed for increased sensitivity in high biomass regions and vegetation enhancement by decoupling the canopy background signal and reducing aerosol and atmospheric influence. Compared to NDVI, it minimizes many of the contamination issues and does not become easily saturated. As a result, it is a strong alternative for forested areas of the planet with large amounts of green vegetation. It is also possible to efficiently estimate plant canopy growth with the physically based plant phenology index (PPI; Jin and Eklundh, 2014), thereby facilitating better vegetation monitoring.

NPP, which is the difference between photosynthesis or GPP and autotrophic respiration over a certain time period (Clark et al., 2001), is a key variable in the biological and chemical carbon cycle and an important ecological indicator of the sustainable development of Earth's ecosystems. As a result, examining NPP variability over time is crucial in assessing global carbon balance and climate change (Potter et al., 2003; Tao and Zhang, 2010). The integration of light use efficiency and process model algorithms has been considered to be a potentially effective means of estimating regional and global NPP using remote sensing data and biophysical/ecological processes (Liu et al., 2001). Despite many achievements in this area in the past few decades, most researchers have focused on the global or national scale. However, this does not take into account that NPP is highly variable over space and time due to

differences or inconsistencies in both anthropogenic and natural factors (Matsushita et al., 2004; Peng et al., 2008).

NPP determines the vegetation atmospheric carbon sequestration and storage rate (Huang et al., 2010; Lu et al., 2009) and has been used to indicate ecosystems decline (Running et al., 2004; Seaquist et al., 2003). Therefore, NPP can be employed both as a baseline and an indicator of forest degradation (Romero-Sanchez and Ponce-Hernandez, 2015), if we assume that canopy cover losses will affect the capacity of the forest to fix carbon, reducing NPP rates.

Tasseled cap transform (TCT) has been employed extensively for the mapping of vegetation and monitoring of changes in land cover and use (Cohen et al., 1995; Collins and Woodcock, 1996; Franklin et al., 2002). The TCT has also been employed as a means of differentiating forest cover densities, assessing the proportion of vegetation to non-vegetation in an area (Gomez et al., 2011; Ahmed et al., 2014), and determining land cover classification (Hermosilla et al., 2022). It both furnishes a means for reducing data volume with very little loss of information and its spectral features are also able to be directly connected with physical land surface parameters (Crist and Cicone, 1984; Crist and Kauth, 1986; Crist et al., 1986). TCT involves converting satellite bands into a set of weighted sums of separate channel readings. One weighted sum measures the 'brightness' of each scene pixel, another represents the degree of pixel 'greenness,' and another the degree of vegetation 'yellowness' or soil 'wetness' (Kauth and Thomas, 1976). The transformation produces a series of orthogonal components–tasseled cap brightness (TCB), tasseled cap greenness (TCG), and TCW. Other components include tasseled cap distance (TCD) and tasseled cap angle (TCA).

The first component, brightness, is a weighted sum of all the bands and is defined in the direction of principal variation in the reflectance of soil, meaning it measures the brightness or total reflectance of soil. The second component, which is greenness, involves the contrast between the NIR and visible bands. The

large amount of scattering that comes from infrared radiation as a result of the cellular structure of green vegetation and also the absorption of visible radiation by plant pigments (e.g., chlorophyll) act together to create high greenness values for target areas that boast high densities of green vegetation; conversely, the flatter reflectance curves that come from soils are expressed in low greenness values. Indices that utilize the contrast between short-wave and near-infrared reflectance, such as TCW, have been some of the more effective means of describing forest vegetation structural differences (Cohen and Goward, 2004). However, since MSS sensors do not have short-wave infrared bands, TCW cannot be computed for pre-1982 Landsat data.

3.1.2 Image fusion

With image fusion, it is possible to combine multisource imagery by means of advanced image processing techniques. Its goal is the integration of data that are both disparate and complementary to enhance the information apparent in the images and interpretation reliability, leading to data that are more accurate (Keys et al., 1990) and useful (Rogers and Wood, 1990). In addition, data that are fused provide for robust operational performance, including reduced ambiguity, as well as improved reliability, confidence, and classification (Rogers and Wood, 1990).

Image fusion that is employed for change detection makes use of the various configurations of the sensor-carrying platforms. Through the combination of the temporal images, we can improve our knowledge of changes that might have taken place in the location under investigation. Temporal image fusion can be applied to images from the same sensor and multiple sensors. Since it is generally not possible to obtain multisensory data simultaneously, data fusion from different sensors usually includes a temporal factor. Mouat et al. (1993) provide a synopsis of image enhancement techniques for change detection with multisensory data (e.g., PCA, image differencing, and ratioing).

Given the fast pace of modern technological advancements, information can now be obtained from multisource images in order to create a high-quality fused image containing both spatial and spectral

information (Susmitha and Pancham, 2009; Huang, 2010). The process of image fusion involves combining multiple images into one image with the greatest amount of information and that does not involve the production of details not present in the original images (Susmitha and Pancham, 2009; Zhang and Blum, 1995). It also results in an image that is more informative and more suitable for visual perception or computer processing (Goshtasby and Nikolov, 2007).

As part of a novel area of research related to statistical/numerical methods, spatiotemporal data fusion has emerged. This method employs statistical or mathematical frameworks in order to create high spatial and high temporal resolution satellite images through the combination of images from different satellite systems with similar spectral and orbital characteristics but different temporal and spatial resolutions.

3.1.2.1 Image fusion techniques (PCA, Wavelet transforms)

A variety of methods exist to perform image fusion. This study examines the following approaches: principal component analysis (PCA) and wavelet transforms—discrete wavelet transforms (DWT) and continuous wavelet transforms (CWT).

PCA

PCA is a statistical analysis method employed in dimension reduction. By means of projecting data from its original space to its eigenspace, it increases the variance and reduces the covariance by keeping the components that correspond to the largest eigenvalues and getting rid of other components. It is easy to implement the linear transformation PCA in applications involving the analysis of a huge amount of data. PCA is commonly employed in both data compression and pattern matching by expressing the data in a manner that highlights the similarities and differences with very little information loss.

The mathematical procedure of PCA involves the transformation of a number of correlated variables into a number of uncorrelated variables, which are termed principal components (PCs). It is used to compute a description of the dataset that is both concise and optimal. The first PC (PC1) accounts for as much of

the data variance as possible (eigenvalue), and every subsequent component accounts for as much of the remaining data variance as possible. In terms of the components, PC1 takes the direction that explains the maximum variance. The next PC (PC2) points to the direction that explains the maximum variance after PC1. The third PC (PC3) is taken in the direction that explains the maximum variance in the subspace perpendicular to the first two principal components, etc.

For image encoding, data compression, enhancement, and fusion; digital change detection; and multi-temporal dimensionality, PCA is very useful. With this statistical technique, it is possible to transform a multivariate dataset composed of intercorrelated variables into a dataset of new and un-correlated linear combinations of the original variables. It also produces a new set of orthogonal axes, which could be used to visualize the distribution of data. By using the standardized PCA, the signal-to-noise ratio can be greatly improved (Singh and Harrison, 1985; Shettigara, 1992).

Wavelet transforms

The wavelet transform, which is quite a relatively new signal analysis tool, furnishes a systematic means for analyzing signals at a variety of resolutions and scales. By employing CWT, it is possible to analyze signals (e.g., multidimensional signals such as image cubes) across a continuum of scales. Signals are analyzed over a discrete set of scales by using DWT. This method enables image decomposition in various types of coefficients preserving the information from the image. These coefficients that are derived from different images can be stitched together to get new coefficients, resulting in the appropriate collection of the information in the original images. After the merging of the coefficients, the resulting fused image is created via inverse discrete wavelets transform, where the information found in the merged coefficients is also preserved. In general, it is possible to accomplish the transform with various fast algorithms and customized hardware, including multiresolution dyadic filter tree implementation.

Various areas of remotely sensed signal processing have benefitted from the application of CWT and DWT. Wavelets have mainly been used for data compression in remote sensing at present (Werness et al., 1994; Luigi-Dragotti et al., 2000). In remote sensing, wavelets can also be applied to and have shown themselves to be robust in areas that include image texture analysis (Fukuda and Hirosawa, 1999), image denoising (Zheng et al., 1999), image edge detection (Simhadri et al., 1998), and data fusion (Nunez et al., 1999).

For multiscale spatial pattern analysis, wavelets are very robust methods of decomposing data into specific components and associating parts of total signal variance with those scales (Csillag and Kabos, 2002). They have been revealed to be more precise in the identification of periodicity than other existing scale-specific techniques (Saunders et al., 2002). As a result, wavelet analysis has been used for analyzing temporal and spatial ecological data (Cazelles et al., 2008), such as examining gap structures in forest canopies (Bradshaw and Spies, 1992), regions of net primary productivity in grasslands (Csillag and Kabos, 2002), and relationships between understory vegetation and the canopy in boreal forests (Kembel and Dale, 2006).

After reviewing and comparing PCA and wavelet transform, PCA was selected as the best method for the humid/semi-arid study area in central and northeastern Ontario, Canada. The PCA technique was applied for the combination of all individual FDIs (1972–2020) into the most informative composite forest degradation indicator (CFDI) and all long-term trajectory maps of 1972–1981, 1984–1998, and 2001–2020 into the most informative disturbance and degradation map of 1972–2020.

3.1.3 Trajectory analysis

3.1.3.1 Trajectory analysis and change detection analysis

Remotely sensed data have been the most important data source for environmental change study over the past few decades and large collections of remote sensing imagery have made it possible to analyze spatio-

temporal patterns of environmental elements and the impact of human activities. The significant amount of optical satellite imagery that is available for free to researchers (e.g., Landsat) (Woodcock et al., 2008) offers opportunities to comprehend and reconstruct the dynamics, especially changes, in forests (Wulder et al., 2008).

Despite the development of many change detection methods (DeVries et al., 2015; Kennedy et al., 2010; Pflugmacher et al., 2012), the challenge of how to reliably map land use change remains. While changes in land use can lead to both conversions and disturbances in land cover, remote sensing generally focuses on the mapping of conversions but not disturbances. However, it is possible that modifications may be more prevalent and meaningful when it comes to forest degradation than conversions. Degradation in the forest context is very likely to be a result of a land-cover modification with its particular duration and degree of intensity.

Temporal trajectory analysis involves analyzing the sequence of changes in detection in every pixel of the image portion of the stack pile of imagery over a time series that is continuous. It has been proven to be especially useful in characterizing the dynamics of land ecosystems owing to the fact that it is efficient in exploiting a multi-temporal sequence of images to determine temporal patterns occurring over several temporal series (Hilker et al., 2009; Pflugmacher et al., 2012). Trajectory-based change detection is automated, requires no screening of non-forested area, and does not need any metric-specific threshold development. Moreover, the method simultaneously provides estimates of discontinuous phenomena (disturbance date and intensity) as well as continuous phenomena (post-disturbance regeneration).

One of the important applications of remote sensing using satellite images is the change detection of objects. Remote sensing change detection identifies differences in intensity, frequency, spatial and temporal extent, as well as stability, rates, and speed between two or more images (Singh, 1989). It can monitor long-term surface condition changes (Wulder et al., 2008) and record their driving forces.

Techniques used in previous change detection studies (Coppin et al., 2004; Lu et al., 2004) did not take full advantage of the interrelationships among many multi-temporal images, and they could be incapable of distinguishing background noise from the subtle or long-term changes expected with climate change (Hicke et al., 2006). To improve signal-to-noise ratio in change detection, methods that consider the signal from multiple images have been reported by Huang et al. (2010) and Wulder et al. (2012).

Using high-resolution optical remote sensing, various change detection techniques have been developed for forest monitoring (in particular clear-cut, regrowth, damage, and disease) (IPCC, 2003; Mikwa et al., 2016). Vegetation dynamics and forest cover changes found with the "greatest change metric," similar disturbance-, or recovery-based metrics try to identify vegetation dynamics that have happened prior to the date of image classification (e.g., Roy et al., 2014). Since land cover composition is linked to its disturbance history, the inclusion of disturbance-related variables should improve land cover classification accuracy as opposed to using single-date spectral variables.

Change detection approaches can be characterized into two broad groups, namely, bi-temporal change detection and temporal trajectory analysis (Coppin et al., 2004). The former measures land cover changes based on a 'two-epoch' timescale (i.e., the comparison between two dates). Even if land cover information is sometimes acquired for more than two epochs, the changes are still measured on the basis of pairs of dates. In contrast to bi-temporal change detection, temporal trajectory analysis is mostly based on low spatial resolution images (e.g., MODIS), which have a high temporal resolution. The trade-off of using these images, however, is the loss of spatial details that makes auto-classification very difficult with low accuracy, so that temporal trajectory analysis is commonly restricted to, for example, vegetation dynamics in large areas or change trajectories of individual land cover classes. Quantitative parameters, such as normalized difference vegetation index (NDVI) or area of given land cover class, are often used as the dependent variables for the establishment of change trajectories.

3.1.3.2 Vegetation time series analysis

Gradual changes are generally representative of the trends that occur in long time series of vegetation data; however, more abrupt changes are also possible. Changes that happen gradually are usually the result of long-term changes in other factors, including land degradation and management, as well as variability in the climate and water availability in different years (Goetz et al., 2005; Kawabata et al., 2001). On the other hand, abrupt changes generally occur because of disturbances, such as fire, flood/drought, urbanization, and insect infestation. In addition, they can be the result of an ecological- or climate-tipping point, wherein a slight disturbance triggers a significant response that leads to a large (and generally irreversible) change, for example an alteration in the regime of the ecosystem (Lenton, 2013; Scheffer et al., 2001).

As a cyclical component, seasonality systematically repeats itself at certain set time intervals (Hill and Lewicki, 2006). Owing to the fact that vegetation seasonality is primarily driven by its phenology, changes that occur in vegetation seasonality can be a reflection of changes in the drivers of plant phenology (e.g., both precipitation and temperature and the interactions among them). Vegetation trends over the long term are greatly affected by external disturbances, but plant phenology can also be impacted and result in changes to seasonality. There are also other seasonality variations (e.g., albedo alterations resulting from variations in snow and moisture) that have an effect on the surface of the land, which can, as a consequence, influence plant seasonality.

In short, various types of changes (e.g., gradual, abrupt, and seasonal) may be represented in the time series of a remotely sensed VI. While the physical causes of the changes may vary, they share underlying statistical properties, such as linearity or nonlinearity, allowing for their detection and quantification with time series analysis techniques.

3.1.4 Techniques and algorithms for trajectory analysis

Changes that occur over the long term, such as in vegetation greenness (Liu et al., 2016), and other indicators of vegetation health, are examined through the use of annual time series. Given the five decades of records, as well as consistency, robust calibration, and free access since the year 2008, Landsat time series are crucial resources for Earth observation data for environmental research (Woodcock et al., 2008; Wulder et al., 2012). Earth observation time series algorithms are employed for the monitoring of changes that occur quickly, including deforestation (Kennedy et al., 2010; Margono et al., 2012) and wildfires (Goetz and Fiske, 2006; Franks et al., 2013). However, Earth observation data are, unfortunately, both structured and unstructured and noisy, and their direct use is generally hard (Eklundh and Jönsson, 2016). When it is considered that Earth observation data are scaling up in quantity to petabytes, it is beyond the ability of common software tools to capture, manage, and process them within a reasonable timeframe (Taylor-Sakyi, 2012; Chen et al., 2012; Camara et al., 2016). Fortunately, a number of time series algorithms for processing Earth observation data are widely utilized for analyzing the forest cover change process, including the following:

(i) Breaks for additive season and trend (BFAST) (Verbesselt et al., 2010) is used for the analysis of seasonal trends.

(ii) Vegetation change tracker (VCT), which was primarily developed for historical forest change processes that have as their basis spectral-temporal properties (Huang et al., 2010).

(iii) TimeSat, which is employed for the monitoring of seasonal trends of land surface processes considering the seasonal parameters (Eklundh and Jönsson, 2016); and

(iv) Landsat-based detection of trends in disturbance and recovery (LandTrendr) which involves the spectral-temporal segmentation of Landsat time series and complex statistical analysis allowing for the extraction of spatial patterns of land cover change magnitude, duration, and year (Kennedy et al., 2018).

Even when they employ the same input data, all time-series change detection algorithms do not produce similar results (Cohen et al., 2017). The implementation of certain algorithms is difficult, and they are more demanding when it comes to their computation requirements. As a result, care must be taken when selecting the optimal algorithm, and researchers should take into consideration the study objective, application domain, and supporting computational platform.

In this study, the BFAST and LandTrendr algorithms, which were reviewed for long-term trajectory analysis of the study area, are briefly described below.

3.1.4.1 BFAST algorithm

BFAST is able to provide the interactive decomposition of time series into several components (trend, seasonal, and noise) with methods that enable the detection and characterization of changes found within the time series (Verbesselt et al., 2010). The BFAST approach is useful for analyzing different types of time series that are formed via imagery from different satellite platforms (such as Landsat and MODIS). It can also be employed in other disciplines examining seasonal and non-seasonal time series, such as climatology, phenology, and hydrology. It is possible to extend the BFAST algorithm to label detected changes with information on the parameters of the fitted piecewise linear models to the trends within changes. The BFAST monitor is an extension of the algorithms inside BFAST, furnishing functionality for the monitoring of disturbances in time series models (with trend/season/regressor terms) (Verbesselt et al., 2012). A variety of models exist for modelling historical behavior that is stable. As a default in regression modelling, a season trend model (with harmonic seasonal pattern) is utilized.

Both the BFAST and BFAST monitor approaches classify individual pixel changes by fitting a harmonic model, testing observations from the monitoring period, and calculating the medium of the residuals for all expected and actual observations in that period (DeVries et al., 2015).

3.1.4.2 LandTrendr algorithm

LandTrendr is a package of routines that help extract information from time series imagery (e.g., Landsat TM, ETM+). Its process employs straight line segments to simplify the temporal trajectory of each Landsat pixel's time series (Kennedy et al., 2010). In the process of label mapping, user-defined rules are imposed for the conversion of the segmentation output into maps highlighting certain processes of vegetation growth or loss. The maps that are produced display the timing and condition at the start of a growth or loss process; this includes the nature of change during the process and the period over which the change took place (Cohen et al., 2010). Kennedy et al. (2010) detailed a segmentation approach that has been shown to be effective in identifying disturbances within forested areas with long time series originating from the Landsat archive (Ahmed et al., 2014a; Frazier et al., 2014; Pflugmacher et al., 2012, 2013). Unfortunately, this procedure requires more processing to be useful, in particular when it comes to ecological applications (Kennedy et al., 2014).

When the goal is the detection of low-magnitude disturbances using Landsat time series, the signal associated with spectral change from the disturbance might be masked by noise connected with normal temporal variation resulting from atmospheric and geometric corrections, variations in the angle of the sun, vegetation phenology, and sensor degradation. Initially, the LandTrendr algorithm was tested by Kennedy et al. (2010) for the detection of forest disturbance and recovery in the US, Pacific Northwest, Oregon, and Washington. The model was able to outperform the bi-temporal change detection that had previously been conducted (Cohen et al., 1998) to track forest disturbances and detect other trends associated with forest regrowth and phenology in the same study area. Kennedy et al. (2010) discovered that different spectral indices varied in their ability to accurately detect subtle disturbance signals. In this case, the NDVI's performance was less effective than that of the NBR or TCW. Many subsequent studies have shown the effectiveness of LandTrendr algorithms for investigating forest disturbance and recovery patterns (e.g., Cohen et al., 2018; Fragal et al., 2016). In addition, trends related to land cover change

resulting from mining based on LandTrendr-Google Earth Engine (GEE) were successfully tracked in South Africa (Dlamini and Xulu, 2019) and Australia (Yang et al., 2018).

With a single pixel from a generated Landsat VI, this algorithm iterates it to identify breakpoints separating periods of durable changes or stability in spectral trajectory, as well as recording the year of the change. Defined by the year and spectral index value, these break points allow for a representation of the spectral history of a particular pixel as a series of vertices that bound line segments. Such segments make it possible for the tracking of a trajectory between years. The algorithm was developed to remain consistent in predictive mapping over time. An annual forest classification in LandTrendr, for instance, will not generally produce inconsistent model outputs (e.g., young, mature, and old-growth conifer classifications) in the same spatial and temporal settings due to small differences in spectral reflectance coming from elements such as the atmosphere, shadows, or cloud cover.

Following the Landsat image archive opening in 2008 (Wulder et al., 2012), time series analysis of Landsat imagery has progressed at a rapid pace leading to the creation of new algorithms and capabilities that facilitate the process of change detection (Zhu, 2017). The LandTrendr algorithm was at the front line of such innovations (Kennedy et al., 2010), employing temporal segmentation of spectral trajectories to distill and extract meaningful change information from the copious amounts of data provided by the Landsat archive (Kennedy et al., 2007). However, this algorithm still has some challenges. First, since the code was written in a proprietary processing language (interactive data language [IDL]), its cost and learning curve have impeded access. Next, its workflow for building time-series image stacks required considerable human investment in image pre-processing and data management before running the algorithm. Finally, the computing time for running the algorithm is excessive, often, several days were required for a single Landsat scene on a typical computer. Taken together, these issues certainly limited the algorithm to users with significant computational resources and expertise in remote sensing. However,

innovations in Earth observation big data analytics and cloud computing systems have altered our ability to handle such big data (Zhu, 2017). Recently, the implementation of the LandTrendr algorithm in the GEE platform fostered opportunities for the online collection of all variable Landsat data and their processing within a cloud-based environment that boast an application programming interface that is user friendly (API) (Kennedy et al., 2018).

As a multi-temporal segmentation tool, LandTrendr can be employed for the generation of trajectory-based spectral time series results generally absent of inter-annual signal noise. Originally, it was implemented in IDL, the proprietary coding language of ENVI^{TM}. Google engineers relayed the IDL code to the open-source GEE coding platform through JavaScript (Kennedy et al., 2018). Implementing LandTrendr through GEE almost eliminates the burdensome data management, time, and image pre-processing aspects associated with the IDL implementation (Kennedy et al., 2018).

For this study, the BFAST and LandTrendr algorithms were reviewed, applied and compared. LandTrendr was selected as the best method for the study area in central and northeastern Ontario, Canada.

The objectives of this chapter are: (i) to identify a suitable set of FDIs derived from multispectral satellite imagery that can be used in the region and then compared with those defined by international standards, (ii) develop an innovative CFDI to examine the relationship with climate variables and future (2050) projection of forest degradation under climate change, and (iii) long-term trajectory analysis using the FDIs, including CFDI, and generating the disturbance and degradation maps for 1972–2020.

3.2 Materials and methods

The methodological workflow describing the steps taken for the computation of the CFDI from individual forest degradation indices, and the subsequent change detection of CFDI through trajectory analysis is shown in figure 3.1 below.

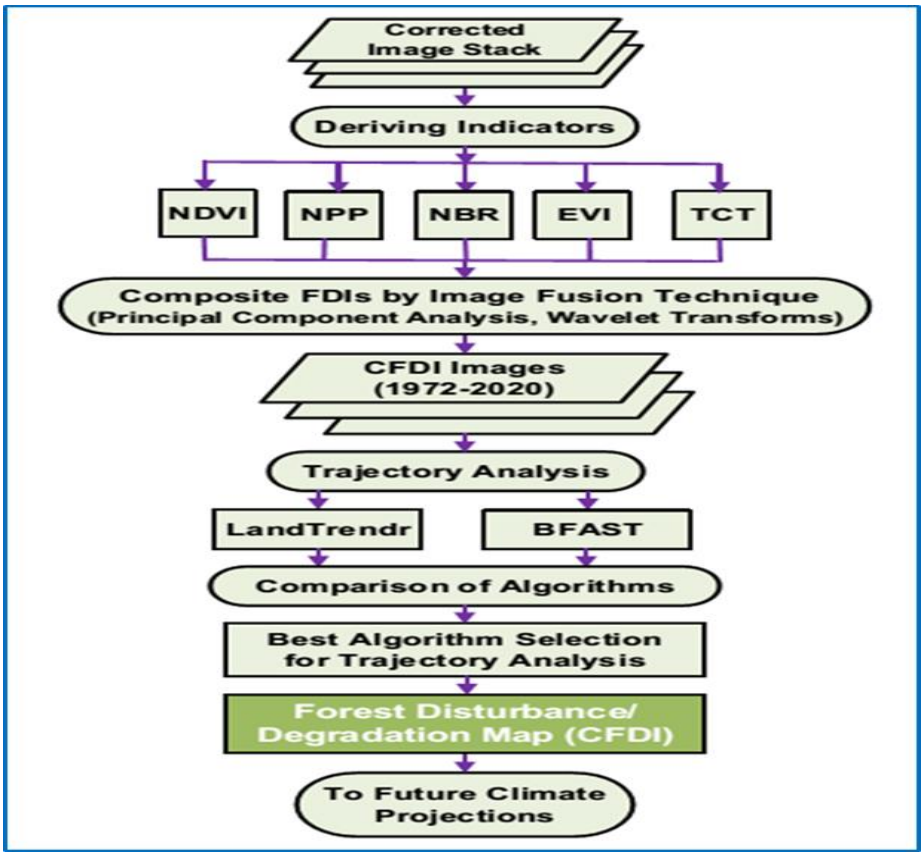


Figure 3. 1: Methodological workflow of CFDI, trajectory analysis, and forest disturbance/degradation trend map

3.2.1 Spectral indices

In total, seven spectral indices—NDVI (Tucker, 1979), NBR (Key and Benson, 2005), EVI (Huete et al., 1997), TCT (TCB, TCG, TCW) (Crist and Cicone, 1984), and NPP (MODIS, 2001–2020)—were calculated from multiple band combinations for unmasked pixels. These were employed as a means of testing the time series trajectory response performance with regards to forest disturbance/degradation. The radiometric transformation takes advantage of the ability to combine information from multiple spectral bands into a single spectral index. The spectral indices are able to form time-series trajectories for the

interpretation of slow or rapid changes. Details and equations for each index are presented in the following sections.

Normalized difference vegetation index (NDVI)

NDVI was developed by Rouse et al. (1974) for the production of a spectral VI that is able to separate green vegetation from background soil brightness with digital data from Landsat MSS. It is expressed as the difference between the NIR and red bands that have been normalized by the sum of those bands. Generally, it is employed as a VI since it maintains the ability to minimize topographic effects while also having the ability to produce a linear measurement scale. NDVI may represent fuel conditions, including both the level of litter and vegetation greenness (Thoha, 1998). The NDVI is particularly successful with extant vegetation spectral indexes because it enables the comparison of inter-annual and seasonal changes that occur in vegetation.

$$\text{NDVI} = \frac{\text{NIR} - \text{R}}{\text{NIR} + \text{R}} \tag{3.1}$$

Normalize Burn Ratio (NBR)

The NBR is used to highlight burned areas in large fire zones (García and Caselles, 1991) and for fire severity assessments. The NBR formula combines both NIR and shortwave-infrared (SWIR) wavelengths, with a high NBR value usually denoting healthy vegetation and a low value indicating bare ground and areas that have recently been burned.

The equation for NBR is similar to that for NDVI, except it uses NIR and SWIR wavelengths. The burn severity map, which highlights burned areas and burn severity, was created with NBR using the following equation:

$$\text{NBR} = \frac{\text{NIR} - \text{SWIR}}{\text{NIR} + \text{SWI R}} \tag{3.2}$$

Enhanced vegetation index (EVI)

EVI, which is an optimized vegetation index, enhances the vegetation signal with greater sensitivity in regions with significant biomass and better monitoring of vegetation by means of a de-coupling of the canopy background signal and a decrease in atmospheric influences. It is linearly correlated with the green Leaf Area Index (LAI) and shows sensitivity to variations in the canopy (Boegh et al., 2002). While it is quite similar to NDVI, it has less sensitivity to atmospheric and background noise. Additionally, its saturation is less than that of NDVI when locations with very dense green vegetation are viewed, especially in LAI regions.

The relevant equation is as follows:

$$\mathrm{EVI} = G \frac{\mathrm{NIR} - \mathrm{Red}}{\mathrm{NIR} + C1 \ X \ \mathrm{Red} - C2 \ X \ \mathrm{Blue} + L} \tag{3.3}$$

Where L is the canopy background adjustment addressing nonlinear, differential NIR, and red radiant transfer occurring through a canopy; and C1 and C2 are the coefficients of the aerosol influences term, which employs the blue band to correct for atmospheric aerosol scattering in the red band. The coefficients that have been adopted in the EVI algorithm are as follows: L=1, C1=6, C2=7.5, and G (gain factor) = 2.5 (Huete et al., 1997). The equation is as follows:

$$\mathrm{EVI} = 2.5 * \frac{\mathrm{NIR} - \mathrm{Red}}{\mathrm{NIR} + 6 * \mathrm{Red} - 7.5 * \mathrm{Blue} + 1} \tag{3.4}$$

Net primary productivity (NPP)

Generally, the estimation of NPP is based on light use efficiency (LUE) theory, which has two assumptions as its foundation: (i) NPP is related to the absorbed photosynthetically active radiation (APAR), with LUE determining the amount of dry matter produced per unit of APAR; and (ii) environmental stress (e.g., low temperature, water shortage) negatively impact LUE (Monteith, 1972; Verstraeten et al., 2006).

NPP can be employed as a direct indicator of the condition and trend of changes in forest ecosystems over time, with a decrease in NPP signaling degradation. NPP is also directly related to the ability of a forest to sequester atmospheric carbon.

NPP can be computed from its component photosynthetically active radiation as follows (Huang et al., 2010; Yuan et al., 2007):

$$NPP = \varepsilon \,.\, fPAR.\, PAR \tag{3.5}$$

Where, PAR is photosynthetically active radiation ($MJm^{-2}month^{-1}$); fPAR represents the fraction of PAR absorbed by the vegetation canopy; ε is the light use efficiency coefficient (g of C/MJ^{-1})) and includes plant respiration costs (Running et al., 2000).

MODIS gap-filled yearly NPP products (MOD17A3HGF-V061) with a spatial pixel resolution of 500m were acquired from the NASA archive (modis.gsfc.nasa.gov). NPP products were obtained based on the difference of the gross primary production (GPP) and the maintenance respiration (MR). The GPP algorithm applies the concept of light use efficiency from both MODIS LAI/FPAR product and also NASA GMAO meteorological data (Running and Zhao, 2015). NPP from 2001–2020 were acquired, 500m data was resampled into 30m, and it was applied in this study to create the CFDI and employ the LandTrendr algorithm to determine the long-term forest disturbance and degradation trends/trajectories as represented by CFDI.

Tasseled cap transformation (TCT)

Tasseled cap indices provide a measure of the brightness, greenness, and wetness (yellowness for MSS data) of a pixel and employ a linear combination of 6 Landsat bands: Landsat TM/ETM+ (bands 1–5 and 7), Landsat 8 (bands 2–7), and MSS bands 1–4 (or 5–8) including two pseudo bands. Each of the TCT indices can be calculated using Tables 3.1, 3.2, and 3.3 and applied to the following equation:

$$TCT_i = Coeff\ i_1 * b_1 + Coeff\ i_2 * b_2 + \ldots\ldots\ldots\ldots + Coeff\ i_n * b_n \quad (3.6)$$

Where TCT_i is the calculated TCB, TCG, and TCW/TCY depending on the coefficients that are employed; $i_1, i_2, \ldots, i_n$ are the coefficients of bands $b_1, b_2, \ldots, b_n$ for Landsat MSS, TM, and ETM+ sensors and bands $b_2, b_3, \ldots, b_n$ (band 2–7) for the Landsat 8 (OLI) sensor.

The reflectance of the surface reflectance data for the tasseled cap coefficients was calculated from a Landsat image on "top of atmosphere (TOA)" calibrated bands, and the coefficients for the TCT for Landsat MSS, Landsat TM, and ETM+ including Landsat 8 (OLI) are shown in Tables 3.1, 3.2, and 3.3 respectively.

Table 3. 1**:** Weight/coefficients of the TCT of Landsat MSS data (according to Kauth and Thomas, 1976)

Component	**Band 1**	**Band 2**	**Band 3**	**Band 4**
Brightness	0.433	0.632	0.586	0.264
Greenness	–0.290	–0.562	0.600	0.491
Yellowness	–0.829	0.522	–0.039	0.194

Table 3. 2**:** Weight/coefficients of the TCT of Landsat TM and ETM+ (according to Crist and Cicone, 1984)

Component	**Band 1**	**Band 2**	**Band 3**	**Band 4**	**Band 5**	**Band 7**
Brightness	0.3037	0.2793	0.4343	0.5585	0.5092	0.1863
Greenness	–0.2848	–0.2435	–0.5436	0.7243	0.0840	–0.1800
Wetness	0.1509	0.1793	0.3299	0.3406	–0.7112	–0.4572

Table 3. 3**:** Weight/coefficients of the TCT of Landsat 8 (according to Baig et al., 2014)

Component	**Band 2**	**Band 3**	**Band 4**	**Band 5**	**Band 6**	**Band 7**
Brightness	0.3029	0.2786	0.4733	0.5599	0.508	0.1872
Greenness	–0.2941	–0.243	–0.5424	0.7276	0.0713	–0.1608
Wetness	0.1511	0.1973	0.3283	0.3407	–0.7117	–0.4559

The reflectance of the surface reflectance data was calculated from Landsat image on "top of atmosphere (TOA)" measurement/calibrated bands. (Tasseled cap coefficients were calculated for TOA reflectance data from the Landsat TM and ETM+ sensors, MSS).

As previously stated, the TCT of the TOA calibrated bands of OLI are computed using a similar method to those employed for images from Landsat 7 (ETM+). It is only the greenness, brightness, and wetness that are determined with the weights presented in Table 3.3. The following equations (3.7, 3.8, and 3.9) are examples of the Landsat 8 OLI sensor.

$$TCB= 0.3029\ b2 + 0.2786\ b3 + 0.4733\ b4 + 0.5599\ b5 + 0.508\ b6 + 0.1872\ b7 \quad (3.7)$$

$$TCG= -0.2941\ b2 - 0.243\ b3 - 0.5424\ b4 + 0.7276\ b5 + 0.0713\ b6 - 0.1608\ b7 \quad (3.8)$$

$$TCW= 0.1511\ b2 + 0.1973\ b3 + 0.3283\ b4 + 0.3407\ b5 - 0.7117\ b6 - 0.4559\ b7 \quad (3.9)$$

3.2.2 Satellite image fusion approaches

Spatiotemporal data fusion approaches

Spatiotemporal data fusion methods were introduced by Gao et al. (2006) as part of their work to produce daily synthetic surface reflectance images at 30m spatial resolution through the combination of both Landsat (30m spatial resolution at 16-day intervals) and MODIS (500m spatial resolution at 8-day intervals) surface reflectance images. This effort aimed at the greatest use possible of imagery data.

A number of other spatiotemporal data fusion techniques have been developed lately (Zhang et al., 2013b; Fu et al., 2013; Weng et al., 2014) and evaluated with different satellite images (e.g., Landsat 5 & 7 and MODIS) (Gao et al., 2006; Fu et al., 2013). The primary goal was the fusion of low-temporal/high-spatial resolution data with high-temporal/low-spatial resolution data in order to improve interpretations in various environmental applications, *inter alia*: (i) surface temperature (Weng et al., 2014), (ii) NDVI (Meng et al., 2013), (iii) surface reflectance (Gao et al., 2006; Zhang 2013), and (iv) evapotranspiration (Anderson et al., 2011). These spatiotemporal data fusion techniques fall into three groups: (i)

spatiotemporal adaptive fusion approaches, (ii) unmixing-based spatiotemporal fusion approaches, and (iii) sparse representation-based spatiotemporal fusion approaches. In this study, spatiotemporal adaptive fusion approaches were employed.

Spatiotemporal adaptive fusion

Mathematical operations were employed by these methods to complete the data fusion. Gao et al. (2006) presented the STARFM that is used to blend surface reflectance data from Landsat and MODIS imagery. This empirical fusion technique employs the spectral and orbital similarities between Landsat and MODIS, combining them for the production of a synthetic Landsat image that has temporal resolution (resampled MODIS 500m yearly data into 30m).

By combining data with different characteristics (varying in spectral, spatial, and temporal resolution), images that are fused can provide better interpretation capabilities and results that are more reliable because they give a more complete view of the observed objects.

In addition to the creation of new images better suited to human visual perception, the goal of image fusion is also object detection and target recognition. Using multisensory data, including both visible and infrared images, has resulted in better recognition rates in various applications (e.g., the automatic recognition of targets) (Zhou and Hecht-Nielsen, 1993). In order for image fusion to succeed, multi-sensor images must be in correct alignment on a pixel-by-pixel basis.

3.2.3 Principal component analysis (PCA)

PCA is a statistical transformation that involves the conversion of intercorrelated multispectral bands into a set of uncorrelated components by means of orthogonal reprojections of the spectral space. A reverse PCA transform is employed to fuse the high-resolution image into the low-resolution multispectral bands (Zhang, 2004; Basaeed et al., 2013). It is possible to calculate the same number of output PC bands as input spectral bands. The variance decreases incrementally with each spectral band, with the largest

percentage of data variance in the first band. As they contain very little variance, most of which is due to noise in the original spectral data, the final PC bands seem noisy. PC bands create more colorful composite images than spectral color composite images owing to the fact that the data is uncorrelated. It is possible for the ENVI software used to analyze these images to accomplish both forward and inverse PC rotations (Richards, 1999 [ENVI^{TM}]).

3.2.4 Long-term trajectories and trends of FDIs and CFDI

In this study, the LandTrendr algorithm was employed to generate the spectral-temporal trajectories with FDIs and the CFDI indicator during the period from 1972–2020. Except for EVI, all the other six FDIs were applied in the LandTrendr algorithm for the assessment of the trajectories and trends of forest disturbance and degradation from 1972–2020 of the LTSS.

LandTrendr

The cloud-based geospatial processing platform from Google Earth Engine^{TM} provides a large number of user-friendly API for the analysis of free and readily available satellite images, the creation of statistics and maps, and the graphical representation through parallel computing of phenomena that have been investigated (Gorelick et al., 2017). With this platform, two elements work together: Google Earth Engine Playground and Google Engine Explorer (Gorelick et al., 2017; Sidhu et al., 2018). LandTrendr, however, involves a set of trajectories and a spectral-temporal segmentation algorithm beneficial for the detection of annual land cover disturbance and recovery. It takes into account the intensity of the land cover disturbance, that is the magnitude of the change, the event duration in years, and the year when the disturbance or recovery takes place (Kennedy et al., 2010). Such detection is conducted with the collection of Landsat time-series data as its foundation.

LandTrendr functions by fitting the spectral-temporal trajectory to each image pixel in an annual time series and engages in tracing the evolution over time of that same pixel in image time series. Since every pixel is able to tell a story, Landsat has the advantage of being able to provide a historical record of a landscape. By making reference to a single band/index, these trajectories can be employed for the identification of a specific land cover type change event in space and time at the pixel level. This algorithm provides information on the change (year detected, duration, and magnitude; see Figure 3.2).

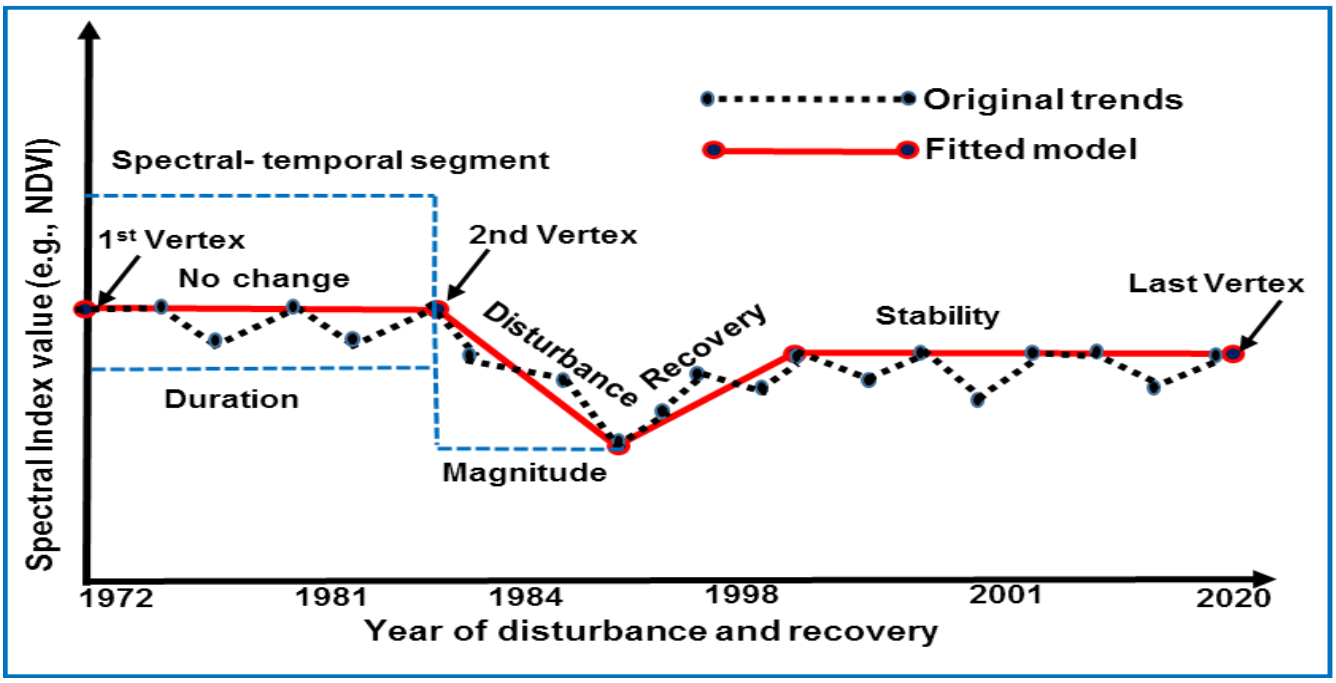


Figure 3. 2: Conceptual/schematic model of the LandTrendr algorithm, original trends and fitting model of the spectral index value (e.g., NDVI) vs year of long-term trajectories

Figure 3.2 illustrates the conceptual model of LandTrendr fitting spectral index (e.g., NDVI) values of spectral-temporal segments for spatiotemporal dynamics of a pixel that has experienced disturbance ("loss"), recovery ("gain"), and stability over the course of 49 years. The first temporal segment, stretching from the first vertex to the second vertex, shows the original model with a slight sequential change. In this illustration, the model is fitted to a no-change event. From the second to third vertices, there was a significant disturbance in the pixel representing an important land cover change. This was followed by a recovery period from the third to fourth vertices. The final land cover changes in this pixel were

characterized by stable interannual variations. This illustration of a conceptual model was adapted from Kennedy et al. (2010).

The LandTrendr program (version 3.0) provided Kennedy et al. (2013) was used with necessary preliminary image processing before radiometric and geometric corrections were undertaken, to create an LTSS compatible to be used with the acquired version of the LandTrendr algorithm.

In order to accomplish the stacking (or collection) of individual Landsat bands for a single image year into one file, the "Layer Stacking" option in ENVI 5.3^{TM} was used (ENVI 5.3, 2015). Both the WGS-84 datum and UTM Zone 18 north projection of the level 1 product generation system (LPGS) processing were retained. Two band-sequential (BSQ) format files were created for the TM/ETM+ images: a six-band stack for the optical bands from 1–5 and 7, as well as one for the thermal band(s). For Landsat 8, six-band layer stacks for band 2 to band 7 were applied.

During the layer stacking procedure, spectral and spatial modifications were necessary for employing the MSS imagery in the assessment. Pixel resampling from 60m to 30m was conducted using a nearest neighbor resampling algorithm to ensure consistent spatial resolution. Subsequently, bands 5 and 7 of the MSS images were interchanged owing to the close approximation of band 7 to the NIR TM/ETM+ band 4 (Bhatt, 2009; Chander et al., 2009). The version of LandTrendr employed during the study was designed to accept only a six-band image stack (for instance, bands 1–5 and 7 for TM/ETM+) and image files with fewer than six bands were rejected (such as four-band MSS data). Two pseudo bands of band 1 and band 6 were inserted into each MSS four-band BSQ file to create the requisite six-band image stack (Table 3.4). These additions were appropriate for this study as the amalgamated image stack was employed only for evaluating NDVI, which assesses band 3 and 4 only. The result was that all LandTrendr preprocessing and data transformations could be applied correctly to the appropriate MSS bands without changing any outputs or results in a significant way.

Table 3. 4**:** Band order for the transformed MSS image stack used in the LandTrendr algorithm

Image stack bands		**True Landsat bands**
Landsat TM/ETM+	**Landsat 8**	**Landsat MSS**
Band 1	Band 2	Band 1/pseudo/false band
Band 2	Band 3	Band 4
Band 3	Band 4	Band 5
Band 4	Band 5	Band 7
Band 5	Band 6	Band 6
Band 7	Band 7	Band 7/pseudo/false band

Note: MSS bands are labelled as bands 4–7 because bands 1–3 were selected for the return beam vidicon (RBV) sensor.

Two preprocessing techniques were used, namely radiometric correction and cloud/cloud shadow masking. These were performed on the LTSS with the LandTrendr standard operating procedures of Nelson et al. (2010) using ENVI 5.3^{TM} and ENVI+IDL 8.5^{TM} software packages (ENVI 5.3, 2015).

The LandTrendr algorithm was employed to generate the spectral-temporal trajectories (for the free version of this algorithm, see http://landtrendr.forestry.oregonstate.edu), thereby producing simplified spectral-temporal trajectories.

In this study, the LandTrendr algorithm detailed by Kennedy et al. (2010) was employed. This algorithm is able to examine each pixel's history, which is termed temporal trajectory (for instance, spectral time series) and consists of yearly index values. Up to six straight-line segments statistically capturing the overall shape of the trajectory are fitted to spectral index values. The potential seven anchor index values are called "vertices" in the fitted temporal trajectory and describe the beginning and end of each linear segment. Through the process of fitting these linear segments, trends or changes that occur in a single spectral index can be better described via the reduction of noise inherent in time series datasets. After the identification of vertices and fitting of linear segments in the target spectral index (here CFDI), the resulting fitted solution can be applied to other spectral indices (e.g., NDVI, NBR, NPP, TCT) and minimizing year-to-year noise across any other spectral index. Here, the terms "temporal segments" and

"temporal segmentation" are employed to refer to the resultant segments and the fitting process, respectively.

The LandTrendr algorithm engages in the decomposition of the trajectory into a series of straight-line segments in order to capture broad features and also sub-trends of the trajectory (Kennedy et al., 2010). This information is necessary to create forest cover change maps or act as predictor variables to forecast future developments. The LandTrendr algorithm is also able to fairly accurately detect minute alterations in an ecosystem, including growth and degradation-related pests/insects (Cohen et al., 2010). LandTrendr temporal segmentation was completed in this study through the running of the algorithm using a fixed set of segmentation parameter values (See Table 2 of Kennedy et al., 2012 or check the parameter set up for this study) (Appendix 3A–G), regardless of the band/index that was employed.

3.3 Results and discussion

3.3.1 FDIs, CFDI, PCA

From 1972–2020, individual FDIs were calculated from the corrected image stack. A total of seven vegetation indices were employed, namely—normalized difference vegetation index (NDVI), net primary productivity (NPP), normalized burn ratio (NBR), enhanced vegetation index (EVI), tasseled cap brightness (TCB), tasseled cap greenness (TCG), and tasseled cap wetness (TCW)—as the indicators selected that best described forest degradation. They were calculated from the satellite images (Landsat and MODIS platforms) using ENVI^{TM} software. All FDI variables were combined into a CFDI for each year, using the image fusion technique reliant on PCA, and the first principal (PC1) was accepted with high explained variability, as shown in Table 3.5.

According to PCA theory, it is possible that the variable responsible for the highest loading could be representative of a PC that is of particular interest (Meng et al., 2009). For instance, (in this study), PC1

consisted of four, six, and seven indicator variables during the periods from 1972–1981, 1984–1998, and 2001–2020, respectively. Table 3.5 shows the explained variability of each PC1 from 1972–2020.

Table 3. 5**:** Explained variability by PC1 of CFDI from 1972–2020

Indicators	Years	Explained Variability	Indicators	Years	Explained Variability
CFDI	1972	99.72%	**CFDI**	2001	96.00%
	1973	99.18%		2002	96.10%
	1975	98.73%		2003	91.67%
	1976	99.64%		2004	96.43%
	1978	99.73%		2005	95.84%
	1979	99.71%		2006	92.66%
	1980	99.75%		2007	93.75%
	1981	99.75%		2009	88.35%
CFDI 1972–1981	**Combined**	**85.01%**		2010	90.19%
CFDI	1984	99.25%		2011	89.17%
	1985	99.40%		2012	91.22%
	1986	99.32%		2013	91.72%
	1987	99.32%		2014	93.08%
	1990	99.33%		2015	92.74%
	1991	99.33%		2016	93.71%
	1992	99.44%		2017	92.99%
	1993	99.46%		2018	93.02%
	1994	99.43%		2019	93.34%
	1995	99.36%		2020	94.35%
	1996	99.31%	**CFDI 2001–2020**	**Combined**	**75.49%**
	1998	99.35%	**NDVI 2001–2020**	**Combined**	**85.55%**
CFDI 1984–1998	**Combined**	**85.69%**	**NPP 2001–2020**	**Combined**	**80.95%**
CFDI combined 1972–2020					**95.32%**

Note: Some years are missing since there is no available and/or usable Landsat image data for those dates.

The CFDI variability from 1972–1981 ranged from 98.73% to 99.75%; from 1984–1998, it ranged from 99.25% to 99.46%. From 2001–2020, the explained variability ranged from 88.35%–96.43%. The explained variability of the combined CFDI of 1972–1981, 1984–1998, and 2001–2020 was 85.01%, 85.69%, and 75.49%, respectively. The explained variability of the combined NDVI and NPP of 2001–2020 was found to be 85.55% and 80.95%, respectively. The explained variability (most informative) of the overall combined (innovative) CFDI for 1972–2020 was 95.32%. According to PCA statistical

interpretation, a combined new indicator (CFDI) is more informative that an individual indicator (FDI); hence, CFDI is advanced as a more informative than an individual FDI.

3.3.2 Remote sensing-derived CFDI maps

The CFDI from 1972–1981 was computed resulting from the multivariate (PCA) as part of the integration of four individual FDIs, namely NDVI, TCB, TCG, and TCW. Figure 3.3(a-b) shows the CFDIs for the initial and final years of 1972 and 1981, respectively.

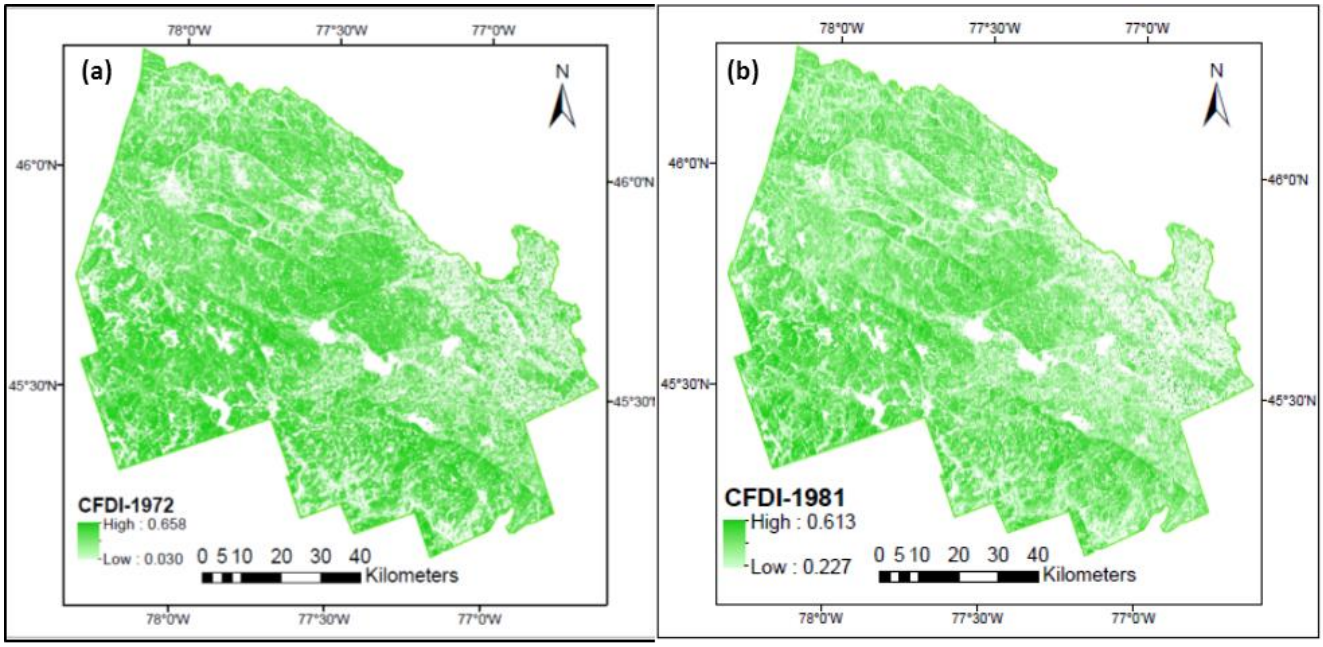


Figure 3. 3 (a–b): a) CFDI of 1972 (left); and b) CFDI of 1981 (right)

It is apparent that the spatial patterns of CFDI from 1972 and 1981 are very similar in their spatial distribution. However, there are changes in the intensity of degradation as shown by the CFDI. There appears to be less degradation in 1972 (0.030) compared to 1981 (0.227), which is to be expected. It is difficult to attribute this to any given cause, either of forest management, ecological phenomena or climate, from the CFDI values alone.

From 1984–1998, the CFDI was computed resulting from the multivariate (PCA) integration of six FDIs, namely NDVI, NBR, EVI, TCB, TCG, and TCW. Figure 3.4 (a-b) presents the CFDIs for 1984 (left) and 1998 (right) (initial and final years, respectively).

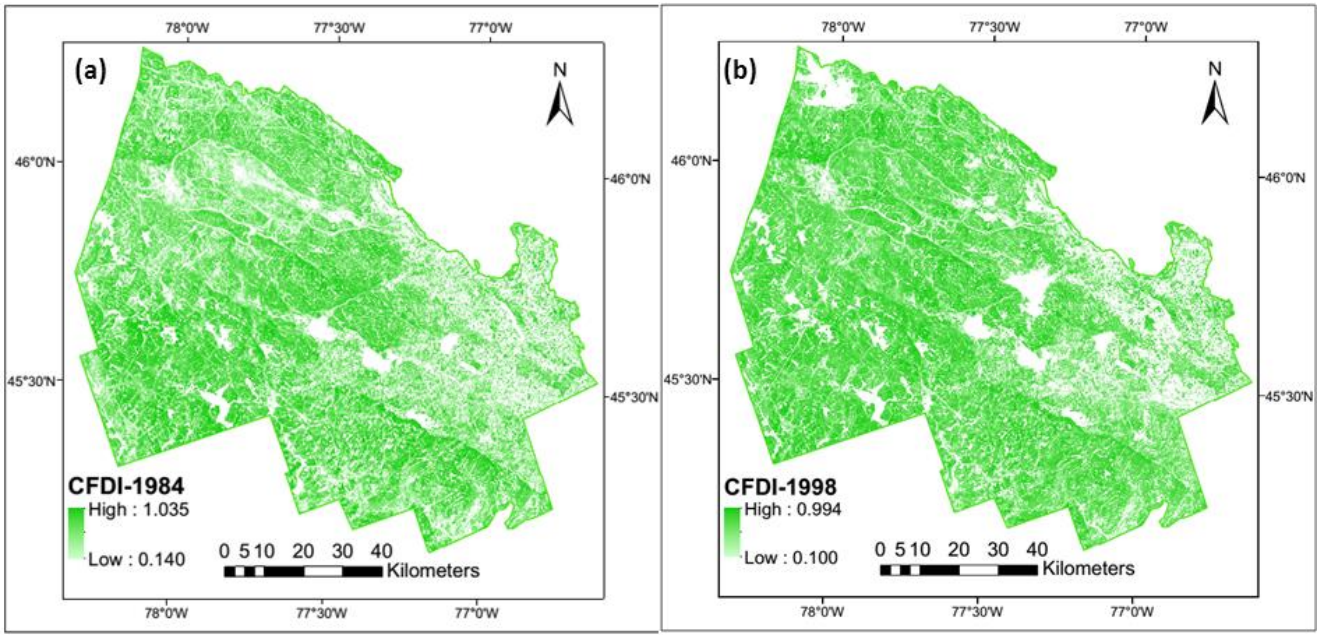


Figure 3. 4 (a-b): a) CFDI of 1984 and b) CFDI of 1998

It is apparent that the spatial patterns of the CFDIs from 1984 and 1998 are not similar. There are changes in the intensity of degradation as shown by the CFDI for 1998. A comparison derived from the visual interpretation of both maps showed that the 1998 CFDI map indicated degraded areas in the northern and middle-eastern part of the study area.

In contrast, the CFDI from 2001–2020 was computed from the multivariate (PCA) integration of seven FDIs: the previous six (NDVI, NBR, EVI, TCB, TCG, and TCW) plus NPP. It was assumed that this would be more explanatory given that NPP would be expected to be more closely related to vegetation health. Figure 3.5 (a-b) shows the CFDIs for the initial (2001) and final year (2020) of the run, respectively.

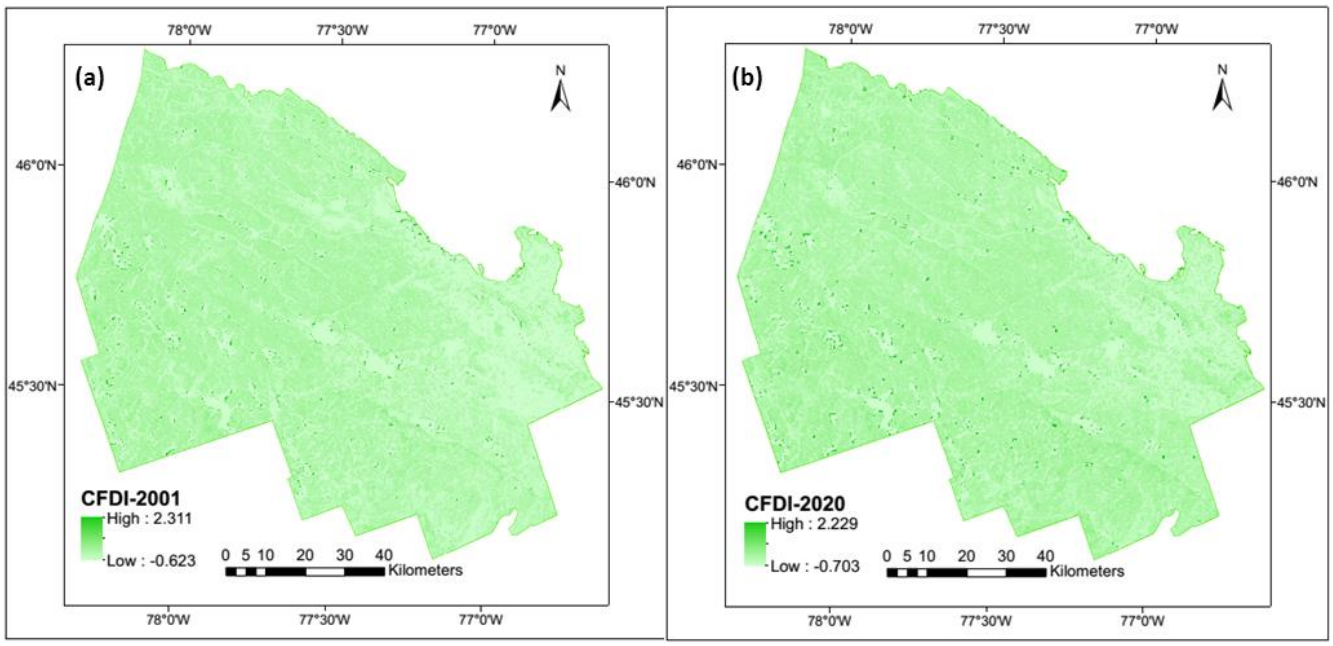


Figure 3. 5 (a–b): a) CFDI for 2001; and b) CFDI for 2020

It is apparent (through visual interpretation) that the spatial patterns of CFDI for the initial year (2001) and final year (2020) are very similar in nature. However, there are changes in the intensity of degradation as shown by the CFDI. The 2001 and 2020 CFDI maps were smoother and finer in resolution than the other CFDI maps (1972–1998). This may be due to the inclusion of the NPP indicator.

3.3.3 Combined CFDI maps of 1972–1981 and 1984–1998

The combined CFDI maps derived from remote sensing are a combination of 1972–1981 and 1984–1998 indicator variables (FDIs). A layer stack was created using the CFDI of each individual year from 1972–1981 before applying the PCA option "Forward PCA rotation new statistics and rotate" in ENVI^{TM} 5.3 software. Another layer stack was created for 1984–1998 using the same procedure.

Overall, the CFDI of 1972–1981 involved four FDIs (NDVI, TCG, TCB, and TCW), and each year was combined for the CFDI using PCA. The CFDI of 1984–1998 involved six FDIs (the four above plus NBR and EVI), and images for each year were combined for the CFDI using PCA.

Finally, the combined CFDI was created with PC1. The results are shown in Figure 3.6 (a and b).

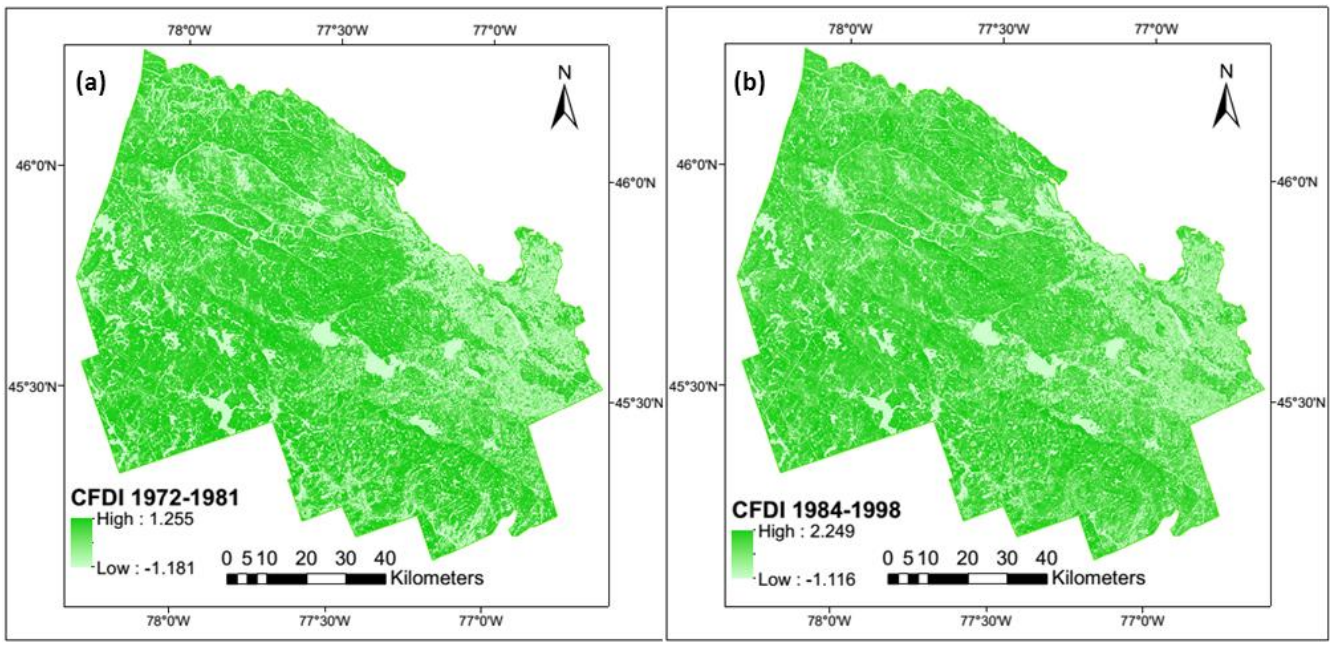


Figure 3. 6 (a-b): a) combination of CFDI from 1972–1981 and b) combination of CFDI from 1984–1998

The explained variability was 85.01% and 85.69% for the combined years of 1972–1981 and 1984–1998, respectively. Through the combination of all variables (FDIs), the most informative in terms of explained variability, was created as the CFDI.

3.3.4 Combined CFDI maps of 2001–2020 and 1972–2020

The composite indicator (CFDI) map values are a combination of 2001–2020 map data and the overall combination of the 1972–2020 CFDI. The CFDI of 2001–2020 involved seven FDIs (NDVI, NPP, NBR, EVI, TCG, TCB, and TCW), with each year of the CFDI combined using the previously mentioned fusion technique. The overall CFDI of 1972–2020 involved all indicators (FDIs) with each year of the CFDI combined using the same fusion technique.

A layer stack was created using the CFDI of each individual year from 2001–2020 using the previously mentioned PCA option in ENVI^{TM}. Additionally, another layer stack was created for 1972–2020 with three CFDI map layers (1972–1981, 1984–1998, and 2001–2020) using the same procedure. Finally, the combined CFDI was created with PC1. These maps are shown in Figure 3.7 (a–b).

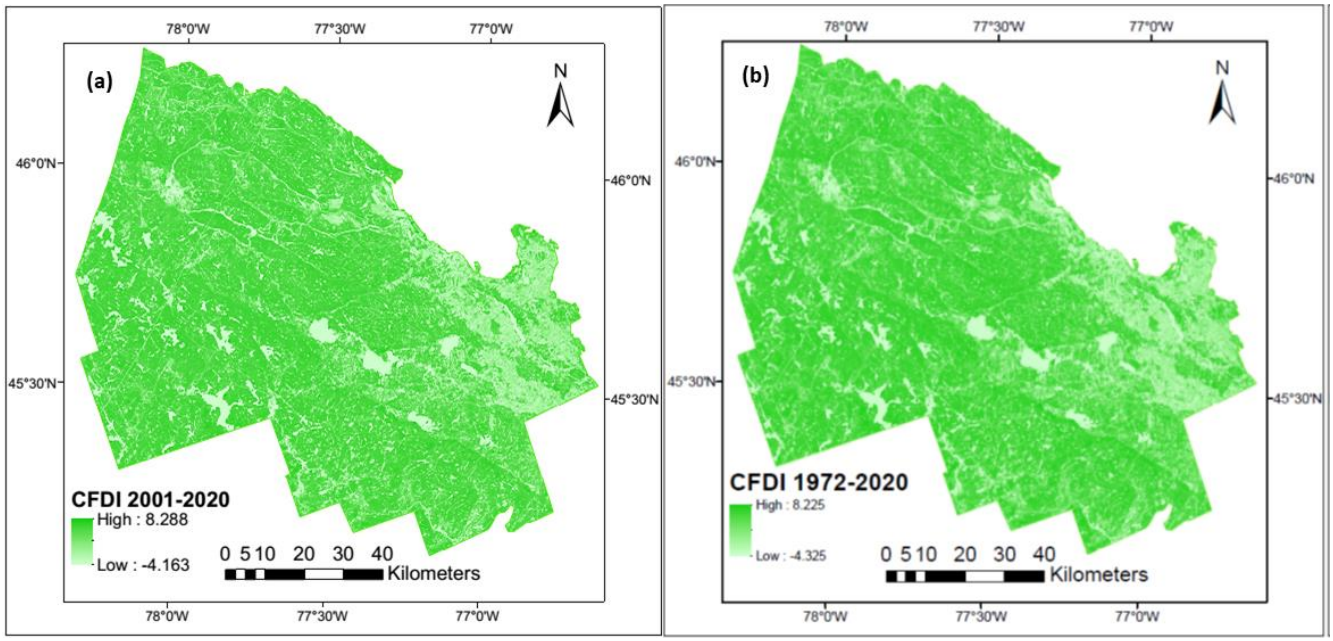


Figure 3. 7 (a-b): a) combination of CFDI from 2001–2020 and b) combination of CFDI from 1972–2020

The explained variability was 75.49% and 95.32% for the combined years of 2001–2020 and 1972–2020, respectively. The CFDI for 1972–2020 is the combined CFDI layer stacks of 1972–1981, 1984–1998, and 2001–2020, with the contribution of all CFDIs providing the most informative CFDI, in terms of explained variability.

3.3.5 Combined maps of NDVI and NPP (2001–2020) from satellite imagery

The combined NDVI and NPP indicator map values are a combination of 2001–2020 individual indicator maps. For the period from 2001–2020. Maps of individual year NDVI and NPP were combined using PCA to create composite NDVI and NPP maps. A layer stack was created using the NDVI of each individual year layer for 2001–2020 before applying the PCA technique. Similarly, another layer stack was created for NPP. Finally, the combined NDVI and NPP map was created by the previously mentioned fusion technique, and the most informative indicator (explained variability) was accepted by PC1. These maps are shown in Figure 3.8 (a–b).

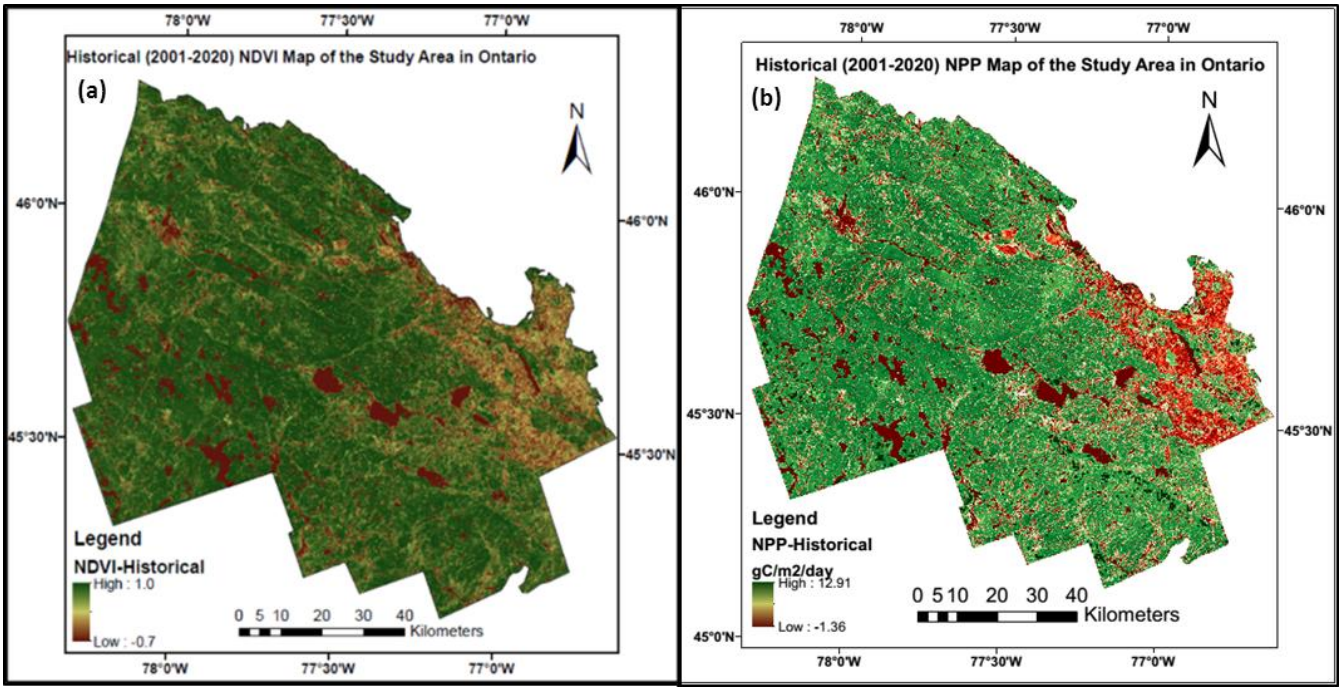


Figure 3. 8 (a–b): a) combination of NDVI from 2001–2020, and b) combination of NPP from 2001–2020

The explained variability for NDVI and NPP were 85.55% and 80.95%, respectively. These combined NDVI and NPP (Figure 3.8a-b) maps are more informative in terms of explained variability (PC1) than individual year maps. A detailed discussion of the combined NDVI and NPP is provided in chapter 4.

3.3.6 Results from the long-term trajectories of forest degradation indicators

As previously mentioned, except for EVI, the trajectory of all the other six FDIs and CFDIs were computed from the LandTrendr algorithm for the assessment of forest disturbance and degradation trajectories from 1972–2020 of the LTSS including MODIS NPP of 2001–2020. The spectral trajectories were calculated and are shown in Figure 3.9 and 3.10. There were many runs of the LandTrendr segmentation algorithm to ascertain the optimal parameters for analysis. For the parameters applied in the final model, see Appendix 3 (A–G).

In this study, ENVI+IDL 8.5 and ENVI Classic 5.3 tools were employed in order to run the LandTrendr algorithm and evaluate the temporal trajectories of the pixels and for the interpretation of the results in the form of output maps. First, LandTrendr was run in “segmentation evaluation mode” to ensure the

existence of no problematic residual clouds or cloud shadows, no images that have been misregistered, or no images that have unpredictably odd phenological signals (LandTrendr Users Guide V3.0, 2013). The batch file that was created for running the LandTrendr evaluation mode is shown in Appendix 3A.

Interpretation of the results of the "segmentation evaluation mode" is complicated. The interpreter must be cautious in determining whether the change is real or an algorithm artifact, meaning the draft disturbance map must be analyzed for suspicious patterns, which then must be looked at for problems, such as issues with cloud mask (LandTrendr Users Guide V3.0, 2013).

Assessment of disturbances involves both visual interpretation and problem-solving for the assessment of the disturbance output maps and anomalous features. If any are found, examination must be made of the core outputs resulting from the segmentation. The images are then subject to "full segmentation mode," and the cloud masks are applied (LandTrendr Users Guide, 2013).

The output files (e.g., "source," "fitted," and "greatest disturbance") of the "segmentation evaluation mode" results have one layer (e.g., in ENVI^{TM} software these are indicated as bands for each year of the time series data, employed as an input image). To interpret the preliminary trajectory maps, three layer bands were used as R, G, and B color maps.

There are three types of output images used to evaluate the segmentation (evaluation mode) runs. The first is the "source" image showing the actual spectral values derived from the pre-processed imagery stack. It is this raw material that is employed by segmentation for the discovery of temporal patterns. The second is the "fitted" image showing how the source image was simplified by the algorithm into straight-line segments; and the third is the "greatest disturbance" image that recorded the disturbance segment in each pixel's time series using the labeling rules (see Appendix 3D: "nbr_label_parameters")

Temporal segmentation is the foundation of LandTrendr. In short, an algorithm is fed a time-series of numbers. This employs statistical rules in order to fit a series of straight lines to the aforementioned

"source" numbers, which then come out as "fitted" numbers. Another set of routines within the algorithm then queries those straight lines while looking for particular sequences or patterns in the data. Generally, the objective is to discover segments consistent with the occurrence of a disturbance, for example, the disturbance that resulted in the most significant change in spectral values (the "greatest disturbance"). While it runs, the algorithm writes files to record the source, fitted, and greatest disturbance numbers of each pixel (LandTrendr v2.0, 2012).

Once the evaluation mode has been run and the user is confident the cloud masks are appropriate and there are no other anomalies in the input images, the batch file can be prepared for running LandTrendr in "full segmentation mode" (see Appendix 3A). To run the LandTrendr algorithm smoothly in "full segmentation mode," in this study, the pre-processed/cleaned satellite images from 1972–2020 were employed.

In LandTrendr, all controlling parameters are contained in separate text files referred to by the main calling sequences, and these files are needed to run the algorithm. There are many parameters (see Appendix 3A–G) that define how LandTrendr segments each pixel. LandTrendr has a range of products based on a variety of combinations of segment label and magnitude of the change, duration, vertex-year, etc. So, a change map can have different change categories, including disturbance types (e.g., fire, clear-cut, insect disturbance), change year, and duration time compared to bivariate change response (simply "change" or "no change") used for traditional change maps.

The trajectory was both fitted and labeled according to the description provided by Kennedy et al. (2010), with the batch file run by LandTrendr presented in Appendix 3A–G. The LandTrendr segmentation algorithm identified periods of time when spectral trajectories were consistently "up" or "down," resulting in segments with vertices and fitted values to be viewed. However, such information is difficult to interpret or summarize in a simple fashion. To make these data more interpretable, this study used an IDL batch file (see Appendix 3A–G) to categorize the segments into "change classes" based on the direction and

magnitude of change, time of occurrence, and duration. For example, capturing the disturbance events, such as fires or landslides (LandTrendr Users Guide, 2013).

Vegetation covers and cover change models are applied for filtering potential segmentation with LandTrendr, and this study ran the filtering batch file with this algorithm (see Appendix 3A–G). Spatial filtering codes are class-specific and added onto the end of the class code. The batch file is presented in Appendix 3A–G.

To examine the long-term trajectory maps (disturbance, recovery, no disturbance, etc.), the change labeling results with 8 layers/bands shown in Table 3.6 were applied. Long-term historical forest degradation trajectory maps "greatest disturbance maps" were created with the LandTrendr algorithm. The standard change label file has 8 layers, as shown in Table 3.6.

Table 3. 6**:** Description of how the bands correspond to the segments

Bands/layers	**Definition of bands/layers**
Band 1	Year of event onset for focal segment
Band 2	Magnitude of event for focal segment
Band 3	Duration of event for focal segment
Band 4	Pre-event vertex value for focal segment
Band 5	Post-event-segment start year
Band 6	Post-event-segment magnitude
Band 7	Post-event-segment duration
Band 8	Post-event-vertex value

The NBR is the most sensitive FDI for capturing disturbance events. A spectral (z) profile was created for the "source" and "fitted" output files to check the "spectral profile" of a single pixel. The "source" image is the NBR value fed to the segmentation algorithm; the "fitted" image is the NBR value that emerges from the segmentation algorithm.

The image of the output results for "greatest disturbance" records the greatest disturbance segment in each pixel's time series, with the labeling rules that are captured in the "label_parameters.txt" (see Appendix 3A–L). Figure 3.9 and 3.10 show an example of spectral trajectories of NBR from the source LTS for 1984–1998 and the final LandTrendr segmentation fitted algorithm of the aforementioned years.

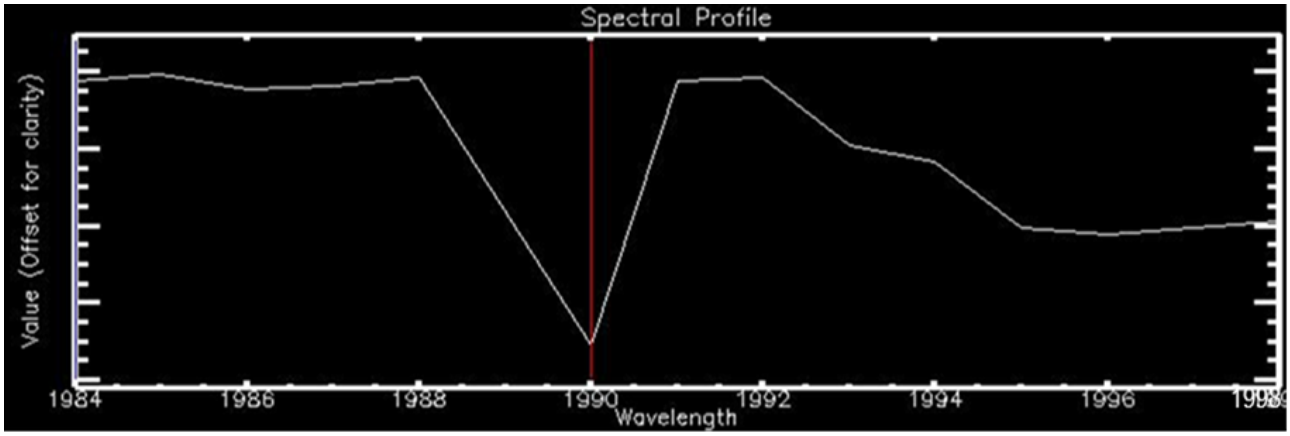


Figure 3. 9: Example spectral profile/trajectories of NBR for the source file (LTSS) outputs (NBR 1984–1998)

The spectral profile of NBR showed in Figure 3.9 indicates greater disturbance and degradation in 1990. The spectral profile for the source image allows for the identification of a very clear disturbance ("loss") or recovery ("gain") trajectory: from 1984–1998, there was a minor disturbance (almost no change); a disturbance occurred in 1990; and recovery following from 1991, with 1992 displaying stability and 1993–1995 indicating the disturbance. The rest of the years (1995–1998) of the trajectory showed a trend of no change.

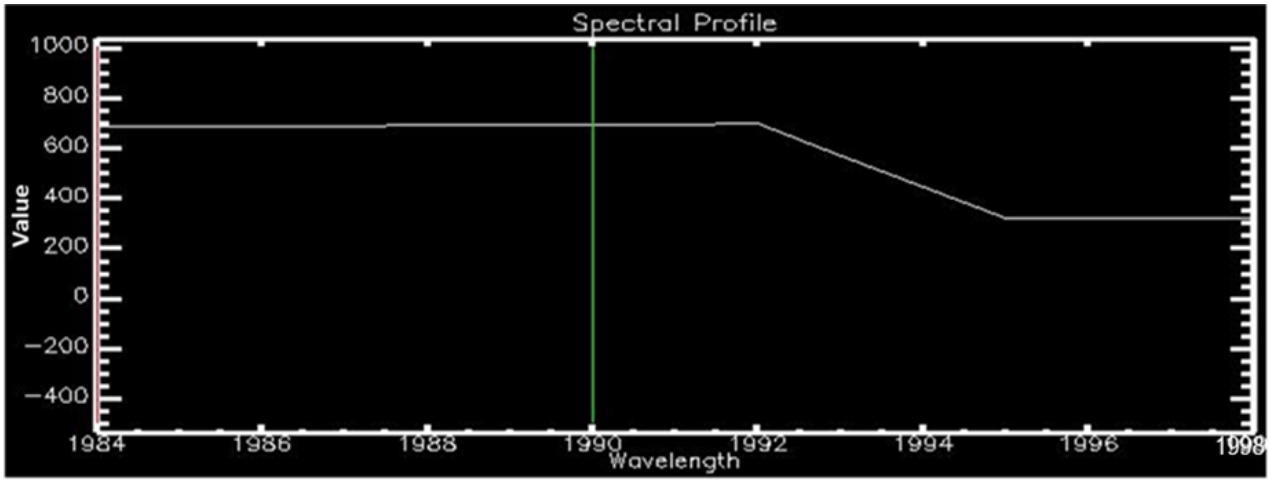


Figure 3. 10: Example spectral profile/trajectories of NBR, the final LandTrendr segmentation fitted outputs for 1984–1998

The spectral profile represents the temporal sequence of the spectral values for a single pixel in the fitted image. There is a big drop in the source image when a large disturbance happens, such as the profile of the trajectory of the NBR indicator from 1984–1998 shown in figure 3.9, which indicates greater disturbance and degradation in 1990.

In this study, the NDVI, NPP, and CFDI indicators were interpreted as being indirectly related to soil moisture availability or "wetness."

3.3.7 Long-term Trajectory of forest disturbance and degradation maps of 1972–2020 (1972–1981, 1984–1998, and 2001–2020)

3.3.7.1 Forest disturbance and degradation maps for 1972–1981

The results of the trajectory analysis from 1972–1981 were employed in the LandTrendr algorithm in order to generate "greatest disturbance maps" of the study area at small and large scales, as shown in Figure 3.11 (a-b). The greatest disturbance map is based on 8 bands and color images, and the following color RGB maps were prepared as b7, b5, and b3.

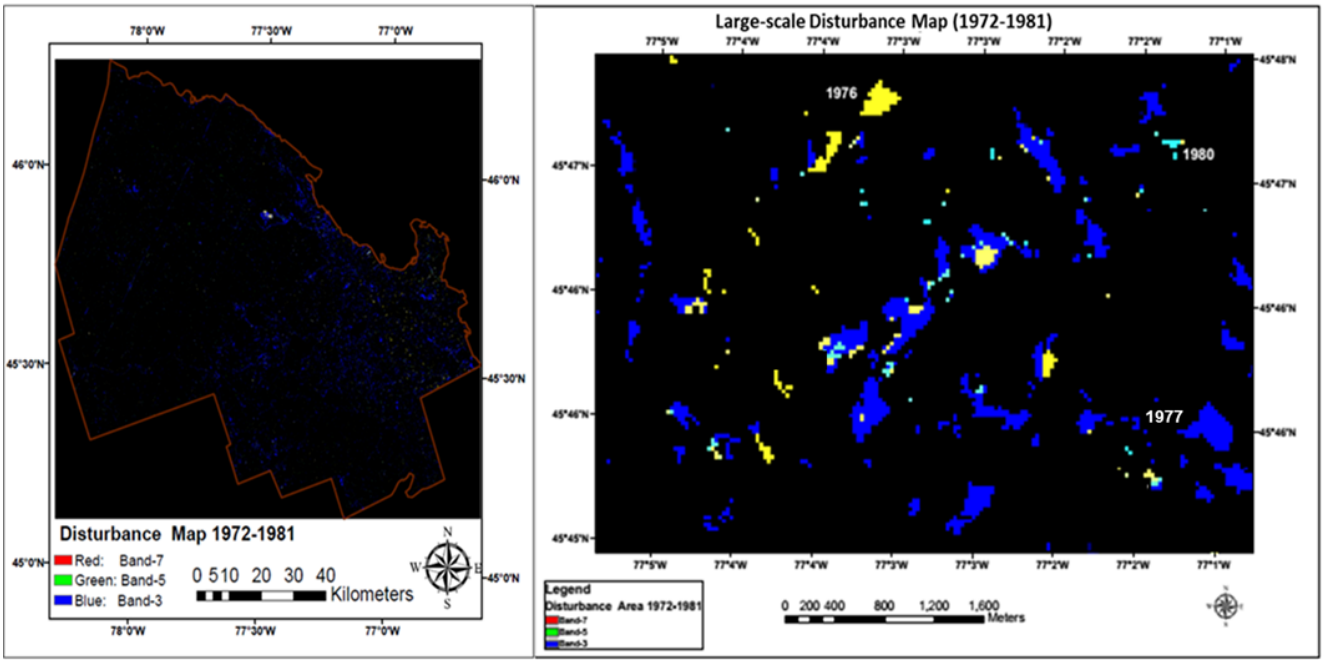


Figure 3. 11 (a-b): Greatest disturbance map (CFDI) of the study area and large-scale map for 1972–1981

Each block of color in Figure 3.11 (a–b) above, indicates the disturbance/degradation or change in landscape during the period from 1972–1981, and the various colors denote different years of disturbance/degradation or change shown on the large-scale map (Figure 3.11a–b) above.

The greatest disturbance map presented all disturbances from 1972–1981 (as an example), whereas the source and fitted maps are presented in 3 layer/band (3 years) combinations (1976, 1977, and 1980 are presented in Figure 3.11b).

3.3.7.2 Forest disturbance and degradation maps for 1984–1998

The results of the trajectory analysis from 1984–1998 were employed in the LandTrendr algorithm in order to generate a “greatest disturbance map” of the study area, which is shown in Figure 3.12.

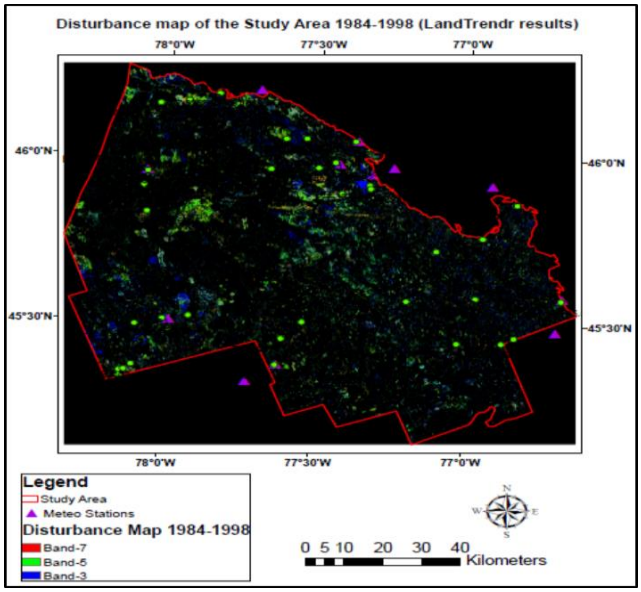


Figure 3. 12: LandTrendr of NBR: Greatest disturbance map of the study area for 1984–1998

Each block of color in Figure 3.12 indicates the disturbance/degradation or change/no change in the landscape during the period from 1984–1998, and the various colors denote different years of disturbance/degradation or change/no change.

3.3.7.3 Forest disturbance and degradation maps for 2001–2020

The results of the trajectory analysis from 2001–2020 and the "greatest disturbance maps" are shown in Figure 3.13 (a-b) (for details, see chapter 4).

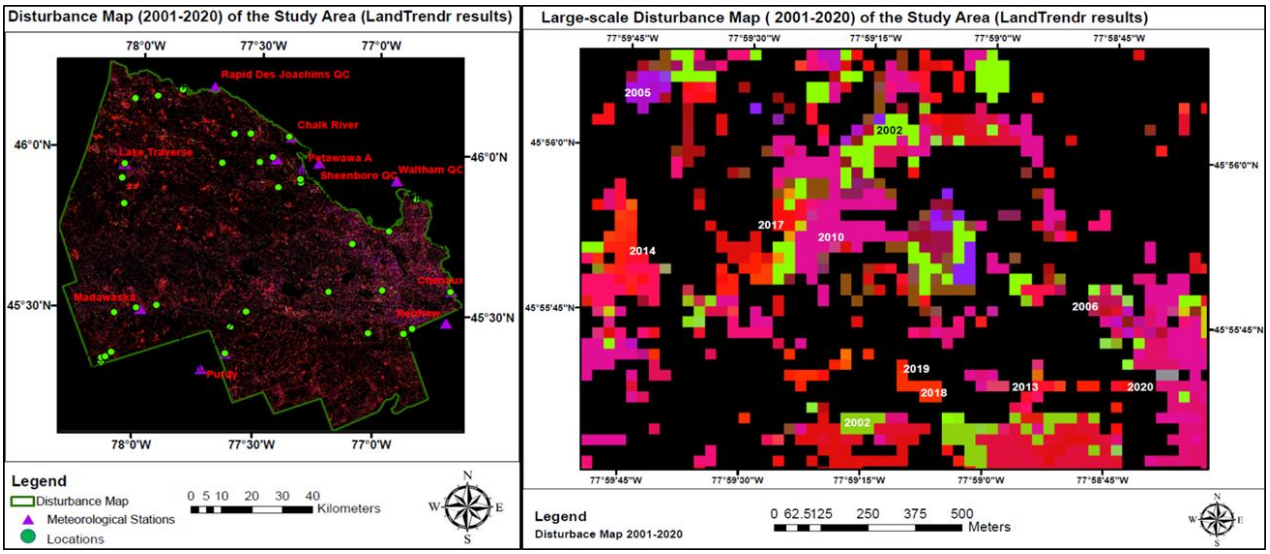


Figure 3. 13 (a-b): (a) Disturbance map (2001–2020, NDVI) of the study area; (b) Large-scale disturbance map

3.3.8 The combined forest disturbance and degradation maps for 1972–2020

Long-term historical forest degradation trajectory maps were created for the periods from 1972–1983, 1984–2000, and 2001–2020 using the LandTrendr algorithm. To ascertain the long-term degradation trend in the selected study area, these three degradation maps were combined into one for 1972–2020 using the image fusion technique (PCA) (Figure 3.14) (a detailed discussion of the combined map is provided in Chapter 4).

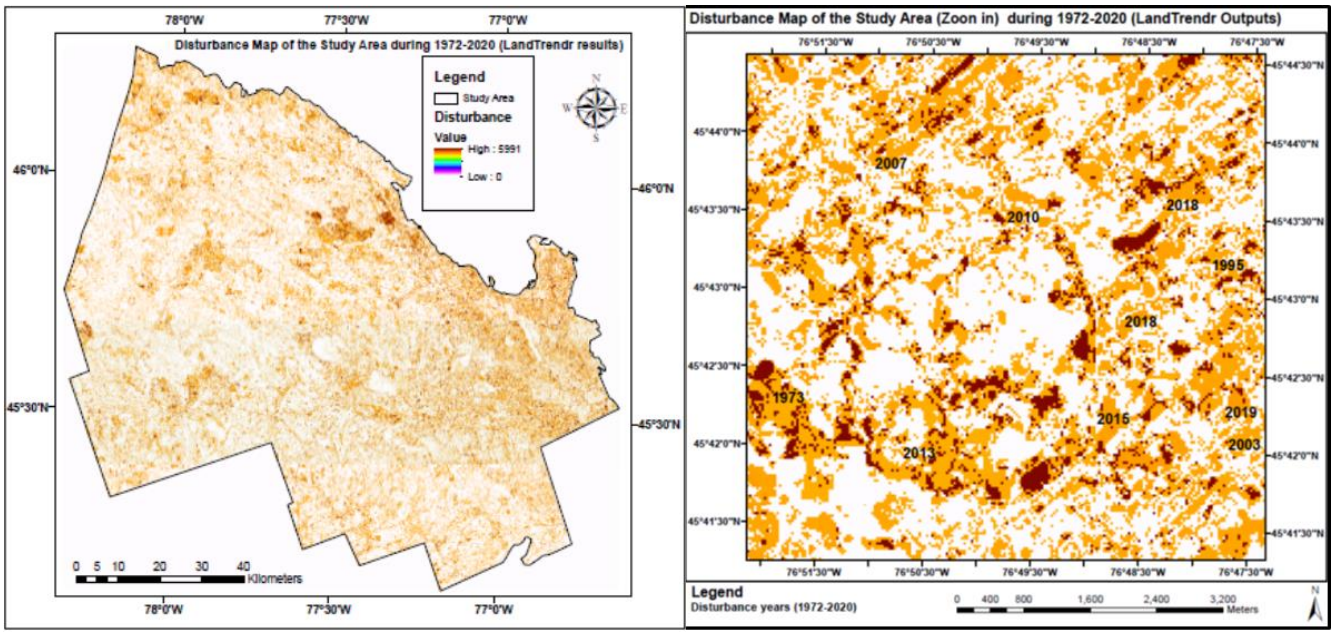


Figure 3. 14 (a-b): a) Disturbance/degradation map (1972–2020) of the study area; b) Large-scale disturbance/degradation map (1972–2020) [Note: PC1 = showing the pixels as values not as year]

By combining the "greatest disturbance" output files of 1972–1981, 1984–1998, and 2001–2020, it was possible to compute a composite disturbance map for the entire time-length of 1972–2020. This was created using PCA for the most informative combined map (PC1) showing the long-term trajectories.

3.4 Discussion

The purpose of monitoring the continuous detection of vegetation dynamics is the production of data and information that are related to the land area, spatial extent, change patterns, and degree of vegetation cover in the study area, which can be useful for forest management decisions. The reconstruction of forest

disturbance histories is necessary for various applications. This study introduced and tested an approach that is efficient and effective for the identification and generation of disturbance histories from the use of the Landsat time series stack. The LandTrendr algorithm was designed to exploit differences in the magnitude and direction of changes that exist in spectral indices (as FDIs in this study), which act in the capacity as "proxies" for the physiological processes that occur after different types of disturbances and degradations in forests. In this work, the capability of the LandTrendr algorithm to achieve the objectives through the use of an illustrative application in the forest cover of the study area was examined.

Through employing the abovementioned algorithm and Landsat imagery, forest disturbance and recovery patterns were detected for each year, and this study presented the forest disturbance and degradation outputs, including raster maps, for the periods from 1972–1981, 1984–1998, and 2001–2020. In this sense, the use of already available and tested algorithms, such as LandTrendr, was found to be convenient and reliable. This study also explored the potential for fusing Landsat and MODIS FDIs into a CFDI that is the most informative indicator of combined forest degradation, by creating a combined map for the long image run from 1972–2020. The resulting disturbance maps and statistics were found to be effective in improving our understanding of the spatial-temporal land cover change trajectories in the study area, over the 49 years from 1972–2020. The results of this thesis showed that Landsat and MODIS derived imagery, despite their differences in spatial and temporal resolutions, can be the basis for the computation of FDIs (NDVI, NPP, NBR, EVI, TCB, TCG, TCW) and that their combination into a CFDI was effective in the identification and mapping of forest degradation.

For the period from 1972–2020, the LandTrendr trajectory analysis was able to clearly identify patterns of forest disturbance, including duration and progress of rehabilitation. However, this study found that maps of forest disturbance derived from the LandTrendr algorithm using the FDIs and CFDI were quite different. This may be due to the information loss that occurs when only PC1 is used as part of the

integration of all FDIs. This study also revealed the critical nature of synergistically employing archived Landsat and MODIS images with trajectory analysis as implemented in the LandTrendr algorithm, for the continuous monitoring of land cover change reconstruction. Moreover, the recently launched Google Earth Engine cloud computing environment provides accessibility to these computing tools. Using combined (1972–1981, 1984–1998, and 2001–2020 disturbance maps) LandTrendr indicators were found to be effective in compensating for the gap values that are commonly seen in single LandTrendr-derived indicators.

This study adopted a combination of LandTrendr-derived indicators to achieve reliable intermediate results. The combination of several indices in characterizing the land surface properties was considered effective. The results were similar to those of Cohen et al. (2017), who emphasized the usefulness of multiple indices in time series analysis instead of employing a single band/index. Additionally, the LandTrendr framework was discovered to be highly customizable and extendable. The indices were combined taking into account PC1 employed to compute the magnitude of the change with filled no-data values. Omitting PCA and simply using LandTrendr with individual FDIs would lead to gaps and discontinuous patches in the maps that could potentially have a negative effect on the interannual land cover change trajectories conducted for the study area.

A wide range of spectral variables of VIs have been employed in Landsat time series analysis based on their enhanced sensitivity to vegetation greenness. The differences that are found in the performance of individual VIs are generally related to their varied sensitivity to external factors, including atmospheric conditions and topography. When choosing the most suitable VI, one must consider the application. In particular, tasseled components are popular in many applications, especially when experts want to combine a multispectral scanner system and data obtained from other sensors. However, more research is

still required into quantifying the extant uncertainties given the varying spectral responses of different sensors in the time series analysis when employing TCT and other indices.

The analysis of the spectral response of vegetation with both remote and proximal sensing could prove to be an efficient and rapid way to examine and define plant health. However, it is still necessary to develop more effective data analysis procedures, especially given the ever-rising amount of data that is being furnished by numerous sensors (e.g., hyperspectral). It is an ongoing challenge to develop methods for automatic fusion of images and interpretation of multi-source, multi-temporal sensor data, especially given the rapid development of various sensor technologies, including VHR optical, SAR, and LiDAR sensors on terrestrial, airborne, and space-borne platforms. Concerning hyperspectral datasets, multivariate statistical techniques may be critical. The combined method applied in this study was based on employing PCA to reduce the dimensionality of the problem with minimum loss of information. PCA was effective in summarizing the redundant radiometric information in a reduced number of indicators of plant or ecosystem status.

In relation to the sensitivity of green vegetation detection, combined individual VIs have led to significant improvements. Various environments provide their own unique variables and complex characteristics, all of which must be taken into account in the selection and use of the different VIs. This means that every single VI has a specific expression of green vegetation, sensitivity for a particular study, and its limitations. VIs, therefore, must be carefully chosen by considering and examining positives and negatives of existing VIs and then combining them for application in a specific environment of interest. VIs can be tailored to specific applications, instrumentation, and platforms to be used in a given study. Since innovations are ongoing in this field, including the development of hyperspectral and multispectral remote sensing technology, new VIs can be created to broaden the research that is possible.

3.5 Conclusions

This study developed a fast and effective approach to monitoring forest disturbance and degradation in the Ontario study area from multiscale, long-term remote sensing data. First, seven indicators (NDVI, NPP, NBR, EVI, TCB, TCG, and TCW) from satellite observation data were selected to describe forest degradation. Then, an approach to forest degradation detection was developed using large stacks of time series Landsat and MODIS data.

All FDI variables were combined into a CFDI for each year, using the image fusion technique reliant on PCA, with PC1 accepted with high explained variability. The explained variability of the combined NDVI and NPP of 2001–2020 were found to be high at 85.55% and 80.95%, respectively. Overall, the innovative CFDI (1972–2020) was the most informative FDI as the combination of three layers of CFDI (1972–1981, 1984–1998, and 2001–2020), and the explained variability of the combined CFDI was 95.32%.

The contribution of all CFDIs provided the most informative CFDI associated with the long-term trajectories. In the spectral profile for the fitted image, there was a very clear disturbance and recovery trajectory: from 1984–1998, there was a minor disturbance; a disturbance occurred in 1990; and recovery following from 1991.

In short, this study demonstrated that exploiting the “big data” from multispectral sensors combining spatial and temporal resolutions has significant potential in monitoring forest degradation at the regional level. The FDIs used in this study, in general, could effectively detect long-term forest disturbance and degradation.

References

Ahmed, O.S., Franklin, S.E., and Wulder, M.A. 2014. Interpretation of forest disturbance using a time series of Landsat imagery and canopy structure from airborne lidar. Can. J. Remote. Sens., 39(6), pp.521–542.

Amiro, B.D., Barr, A.G., Barr, J.G., Black, T.A., Bracho, R., Brown, M., Chen, J. Clark, K.L., Davis, K.J., Desai, A.R., Dore, S., Engel, V., Fuentes, J.D., Goldstein, A.H., Goulden, M.L., Kolb, T.E., Lavigne, M.B., Law, B.E., Margolis, H.A., Martin, T., McCaughey, J.H., Mission, L., Montes-Helu, M.,

Noorments, A., Randerson, J.T., Starr, G., and Xiao, J. 2010. Ecosystem carbon dioxide fluxes after disturbance in forests of North America. Journal of Geophysical Research, 115(G4), pp.1–13.

Anderson, M.C., Hain, C., Wardlow, B., Pimstein, A., Mecikalski, J.R., and Kustas, W.P. 2011. Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States. Journal of Climate, 24, pp.2025–2044.

Asner, G.P. 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sensing of Environment, 64, pp.234–253.

Avery, T.E., and Berlin, G.L. 1992. Fundamentals of remote sensing and airphoto interpretation. Upper Saddle River, NJ: Prentice Hall.

Basaeed, E., Bhaskar, H., and Al-Mualla, E.M. 2013. Comparative analysis of pan-sharpening techniques on DubaiSat-1 images. In, Proceedings of the 16th International Conference on Information Fusion, Istanbul, Turkey, 9–12 July 2013.

Bhatt, R. 2009. Consistent radiometric calibration of Landsat-1 through -5 MSS sensors using pseudo-invariant calibration sites. Department of Engineering: South Dakota State University.

Birkett, C. 1998. Global wetland distribution and functional characterization: trace gases and the hydrologic cycle. In, Santa Barbara, C.A., Dork, S., and John, M. (eds), Joint GAIM-DIS-BAHC-IGAC-LUCC workshop.

Boegh, E., Soegaard, H., Broge, N., Hasager, C.B., Jensen, N.O., Schelde, K., and Thomsen, A. 2002. Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture. Remote Sens. Environ., 81, pp.179–193.

Bonan, G.B. 2008a. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. Science, 320, pp.1444–1449.

Braaten, J. 2012. Processing instructions for LandTrendr v2.0.

Bradley, B.A., Jacob, R.W., Hermance, J.F., and Mustard, J.F. 2007. A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. Remote Sens. Environ., 106, pp.137–145.

Bradshaw, R.H.W., and Spies, T. 1992. Characterizing canopy gap structure in forests using wavelet analysis. J Ecol., 80, pp.205– 215.

Brandt, J.P., Flannigan, M.D., Maynard, D.G., Thompson, I.D., and Volney, W.J.A. 2013. An introduction to Canada's boreal zone: ecosystem processes, health, sustainability, and environmental issues 1. Environmental Reviews, 21(4), pp.207–226.

Brewer, K.C., Winne, J.C., Redmond, R.L., Opitz, D.W., and Mangrich, M.V. 2005. Classifying and mapping wildfire severity: a comparison of methods. Photogrammetric Engineering and Remote Sensing, 71(11), pp.1311−1320.

Brooks, E.B., Wynne, R.H., Thomas, V.A., Blinn, C.E., and Coulston, J.W. 2014. On-the-fly massively multitemporal change detection using statistical quality control charts and Landsat data. IEEE Trans. Geosci. Remote Sens., 52, pp.3316–3332.

Camara, G., Assis, L.F., Ribeiro, G., Ferreira, K.R., Llapa, E., and Vinhas, L. 2016. Big Earth observation data analytics: matching requirements to system architectures. In, Proceedings of the 5th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data, San Francisco, CA, USA, 31 October 2016; pp. 1–6.

Campbell, J.B. 1996. Introduction to Remote Sensing. The Guilford Press.

Cazelles, B., Chavez, M., Berteaux, D., Me´nard, F., Vik, J.O., Jenouvrier, S., and Stenseth, N. 2008. Wavelet analysis of ecological time series. Oecologia, 156, pp.287–304.

Chander, G., Markham, B.L., and Helder, D.L. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sensing of Environment, 113, pp.893–903.

Chapin, F.S. 2005. Role of land-surface changes in Arctic summer warming. Science, 310, pp.657–660.

Chen, W., Li, J., Zhang, Y., Zhou, F., Koehler, K., Leblanc, S., Fraser, R., Olthof, I., and Wang, J. 2009b. Relating biomass and leaf area index to non-destructive measurements in order to monitor changes in Arctic vegetation. Arctic, 62, pp.281.

Chen, Z., Chen, N., Yang, C., and Di, L. 2012. Cloud computing enabled web processing service for Earth observation data processing. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 5, pp.1637–1649.

Clark, D. et al. 2001. NPP in tropical forests: an evaluation and synthesis of existing field data. Ecol. Appl., 11, pp.371–384.

Clark, K.L., Skowronski, N., and Hom, J. 2010. Invasive insects impact forest carbon dynamics. Global Change Biology, 16, pp.88–101.

Cocke, A.E., Fulé, P.Z., and Crouse, J.E. 2005. Comparison of burn severity assessments using Differenced Normalized Burn Ratio and ground data. International Journal of Wildland Fire, 14, pp.189−198.

Cohen, W.B., and Goward, S.N. 2004. Landsat's role in ecological applications of remote sensing. Bioscience, 54, pp.535–545.

Cohen, W.B., Fiorella, M., Gray, J., Helmer, E., and Anderson, K. 1998. An efficient and accurate method for mapping forest clearcuts in the Pacific Northwest using Landsat imagery. Photogram Eng. Remote Sensing, 64, pp.293–300.

Cohen, W.B., Healey, S.P., Yang, Z., Stehman, S.V., Brewer, C.K., Brooks, E.B., Gorelick, N., Huang, C., Hughes, M.J., and Kennedy, R.E. 2017. How similar are forest disturbance maps derived from different Landsat time series algorithms? Forests, 8, pp.98.

Cohen, W.B., Spies, T.A., and Fiorella, M. 1995. Estimating the age and structure of forests in a multi-ownership landscape of western Oregon, U.S.A. International Journal of Remote Sensing, 16, pp.721–746.

Cohen, W.B., Yang, Z., and Kennedy, R. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — Tools for calibration and validation. Remote Sens. Environ., 114, pp.2911–2924.

Cohen, W.B., Yang, Z., Healey, S.P., Kennedy, R.E., and Gorelick, N. 2018. A LandTrendr multispectral ensemble for forest disturbance detection. Remote Sens. Environ., 205, pp.131–140.

Collins, J.B., and Woodcock, C.E. 1996. An assessment of several linear change detection techniques for mapping forest mortality using multitemporal Landsat TM data. Remote Sensing of Environment, 56, pp.66–77.

Coppin, P., Lambin, E., Jonckheere, I., Nackaerts, K., and Muys, B. 2004. Digital change detection methods in ecosystem monitoring: A review. International Journal of Remote Sensing, 25, pp. 1565−1596.

Crist, E.P. 1985. A TM Tasseled Cap equivalent transformation for reflectance factor data. Remote Sensing of Environment, 17(3), pp.301–306.

Crist, E.P., and Cicone, R.C. 1984. A physically-based transformation of thematic mapper data—the TM tasseled cap. IEEE Transactions on Geoscience and Remote Sensing, 22(3), pp.256–263.

Crist, E.P., and Kauth, R. J. 1986. The tasseled cap de-mystified. Photogrammetric Engineering and Remote Sensing, 52(1), pp. 81–86.

Crist, E.P., Lauren, R., and Cicone, R.C. 1986. Vegetation and soils information contained in transformed thematic mapper data. Proceedings of IGARSS' 86 Symposium, pp. 1465–1470.

Csillag, F., and Kabos, S. 2002. Wavelets, boundaries, and the spatial analysis of landscape pattern. Ecoscience, 9, pp.177–190.

De Groof, H., Sowter, A., and Sieber, A.J. 1992. The TREES project: creating an environment for the routine classification and image processing of ERS-1 SAR imagery. International Space Year: Space Remote Sensing, Proceedings 12th Annual IEEE International Geoscience and Remote Sensing Symposium (IGARSS '92), held in Houston, U.S.A., edited by R. Williamson with collaboration of T. Stein (New York), pp. 1343–1345.

DeVries, B., Verbesselt, J., Kooistra, L., and Herold, M. 2015. Robust monitoring of smallscale forest disturbances in a tropical mountain forest using Landsat time series. Remote Sens. Environ., 161, 107–121.

Dlamini, L.Z., and Xulu, S. 2019. Monitoring mining disturbance and restoration over RBM site in South Africa using LandTrendr Algorithm and Landsat data. Sustainability, 11, pp.6916.

Ehlers, M., Klonus, S., Johan Åstrand, P., and Rosso, P. 2010. Multi-sensor image fusion for pansharpening in remote sensing. International Journal of Image and Data Fusion, 1, pp.25–45.

Eklundh, L., and Jönsson, P. 2016. TIMESAT for processing time-series data from satellite sensors for land surface monitoring. In Multitemporal Remote Sensing; Springer: Berlin/Heidelberg, Germany, pp. 177–194.

ENVI (Environment for Visualizing Images) version 5.0. Exelis Visual Information Solutions, Boulder, CO, USA. http://www.harrisgeospatial.com/ProductsandSolutions/GeospatialProducts/ENVI.

ENVI version 5.3. 2015. Exelis Visual Information Solutions, Boulder, CO, USA. http://www.harrisgeospatial.com/ProductsandSolutions/GeospatialProducts/ENVI.

ENVI. 2016. Spectral Indices. Harris Geospatial. Exelis VIS Docs Center. https://www.harrisgeospatial.com/docs/SpectralIndices.html.

Fensholt, R., et al. 2013. Assessing land degradation/recovery in the African Sahel from long-term Earth observation based primary productivity and precipitation relationships. Remote Sens., 5(2), pp.664–686.

Fiorella, M., and Ripple, W.J. 1993b. Determining successional stage of temperate coniferous forests with Landsat satellite data. Photogrammetric Engineering and Remote Sensing, 59, pp.239–246.

Flannigan, M.D., Logan, K.A., Amiro, B.D., Skinner, W.R., and Stocks, B.J. 2005. Future area burned in Canada. Climatic Change, 72(1–2), pp.1–16.

Food and Agriculture Organization of the United Nations (FAO). 2015. Global Forest Resources Assessment 2015, pp.253.

Fragal, E.H. 2015. Reconstrução histórica de mudanças na cobertura florestal em várzeas do baixo Amazonas utilizando o algoritmo LandTrendr. Dissertação de Mestrado, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, São Paulo.

Fragal, E.H., Silva, T.S.F., and de Moraes Novo, E.M.L. 2016. Reconstructing historical forest cover change in the Lower Amazon floodplains using the LandTrendr algorithm. Acta Amazon, 46, pp.13–24.

Franklin, S.E., Lavigne, M.B., Wulder, M.A., and McCaffrey, T.M. 2002. Large-area forest structure change detection: an example. Canadian Journal of Remote Sensing, 28(4), pp.588–592.

Franks, S., Masek, J.G., and Turner, M.G. 2013. Monitoring forest regrowth following large scale fire using satellite data–A case study of Yellowstone National Park, USA. Eur. J. Remote Sens., 46, pp.551–569.

Frazier, R.J., Coops, N.C., Wulder, M.A., and Kennedy, R. 2014. Characterization of aboveground biomass in an unmanaged boreal forest using Landsat temporal segmentation metrics. ISPRS J. Photogramm. Remote Sens., 92, pp.137–146.

Fu, D., Chen, B., Wang, J., Zhu, X., and Hilker, T. 2013. An improved image fusion approach based on enhanced spatial and temporal the adaptive reflectance fusion model. Remote Sensing, 5, pp.6346–6360.

Fukuda, S., and Hirosawa, H. 1999. A wavelet-based texture feature set applied to classification of multifrequency polarimetric SAR images. IEEE Trans. Geosci. Remote Sensing, 37, pp. 2282–2286.

Gao, F., Masek, J., Schwaller, M., and Hall, F. 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. IEEE Transactions on Geoscience and Remote Sensing, 44, pp.2207–2218.

García, M.J.L., and Caselles, V. 1991. Mapping burns and natural reforestation using thematic mapper data. Geocarto Int., 6, pp.31–37.

Gauthier, S., Bernier, P., Burton, P.J., Edwards, J., Isaac, K., Isabel, N., Jayen, K., Le Goff, H., and Nelson, E.A. 2014. Climate change vulnerability and adaptation in the managed Canadian boreal forest. Environmental Reviews, 22(3), pp.256–285.

Gerylo, G.R., et al. 2002. Empirical relations between Landsat TM spectral response and forest stands near fort Simpson, Northwest Territories, Canada. Can. J. Remote Sens., 28(1), pp.68–79.

Giuliani, G., Chatenoux, B., Piller, T., Moser, F., and Lacroix, P. 2020. Data Cube on Demand (DCoD): Generating an Earth observation Data Cube anywhere in the world. Int. J. Appl. Earth Observ., 87, pp.102035.

Gizachew, B., Solberg, S., Næsset, E., et al. 2016. Mapping and estimating the total living biomass and carbon in low-biomass woodlands using Landsat 8 CDR data. Carbon Balance and Management, 11(13).

Goetz, S.J., Bunn, A.G., Fiske, G.J., and Houghton, R. 2005. Satellite-observed photosynthetic trends across boreal North America associated with climate and fire disturbance. Proceedings of the National Academy of Sciences of the United States of America, 102, pp.13521–13525.

Goetz, S.J., Fiske, G.J., and Bunn, A.G. 2006. Using satellite time-series data sets to analyze fire disturbance and forest recovery across Canada. Remote Sens. Environ., 101, pp.352–365.

Gomes, V.C., Queiroz, G.R., and Ferreira, K.R. 2020. An overview of platforms for big Earth observation data management and analysis. Remote Sens., 12, pp.1253.

Gomez, C., White, J.C., and Wulder, M.A. 2011. Characterizing the state and processes of change in a dynamic forest environment using hierarchical spatio-temporal segmentation. Remote Sens. Environ., 115, pp.1665–1679.

Goodale, C.L., Apps, M.J., Birdsey, R.A., Field, C.B., Heath, L.S., Houghton, R.A., Jenkins, J.C., Kohlmaier, G.H., Kurz, W., Liu, S., Nabuurs, G.-J., Nilsson, S., and Shvidenko, A.Z. 2002. Forest carbon sinks in the Northern Hemisphere. Ecological Applications, 12, pp.891–899.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ., 202, pp.18–27.

Goshtasby, A., and Nikolov, S. 2007. Image fusion: advances in the state of the art. Information Fusion, 8, pp.114–118.

Haberl, H., et al. 2007. Quantifying and mapping the human appropriation of net primary production in Earth's terrestrial ecosystems. Proc. Natl. Acad. Sci., 104(31), pp.12942–12947.

Hall, R.J., Skakun, R.S, Arsenault, E.J., and Case, B.S. 2006. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. For. Ecol. Manage., 225(1–3), pp.378–90.

Hamunyela, E. 2017. Space-time monitoring of tropical forest changes using observations from multiple satellites. Ph.D. Thesis, Wageningen University & Research, Laboratory of Geo-Information Science and Remote Sensing, Wageningen, The Netherlands.

Hansen, M.C., et al. 2013. High-resolution global maps of 21st-century forest cover change. Science, 342, pp.850–853.

Hermosilla, T., Wulder, M.A., White, J.C., and Coops, N.C. 2022. Land cover classification in an era of big and open data: optimizing localized implementation and training data selection to improve mapping outcomes. Remote Sens. Environ., 268, pp.112780.

Hicke, J.A., Logan, J.A., Powell, J. Ojima, D.S. 2006. Changing temperatures influence suitability for modeled mountain pine beetle (Dendroctonus ponderosae) outbreaks in the western United States. Journal of Geophysical Research, Biogeosciences, 11, pp. G02019.

Hicke, J.A., Allen, C.D., Desai, A.R., Dietze, M.C., Hall, R.J., Hogg, E.H., Kashian, D.M., Moore, D., Raffa, K.F., Sturrock, R.N., and Vogelmann, J. 2012. Effects of biotic disturbances on forest carbon cycling in the United States and Canada. Global Change Biology, 18, pp.7–34.

Hilker, T., Wulder, M.A., Coops, N.C., Seitz, N., White, J.C., Gao, F., Masek, J.G., and Stenhouse, G. 2009. Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. Remote Sens. Environ., 113, pp.1988–1999.

Hill, T., and Lewicki, P. 2006. Statistics: methods and applications: a comprehensive reference for science, industry, and data mining. StatSoft, Inc.

Huang, C., Goward, S.N., Masek, J.G., Thomas, N., Zhu, Z., and Vogelmann, J.E. 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. Remote Sens. Environ., 114, pp.183–198.

Huang, C., Wylie, B., Yang, L., Homer, C., and Zylstra, G. 2002. Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. International Journal of Remote Sensing, 23, pp.1741–1748.

Huang, S.G. 2010. Wavelet for image fusion. Graduate Institute of Communication Engineering and Department of Electrical Engineering, National Taiwan University.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., and Ferreira, L.G. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ., 83, pp.195–213.

Huete, A., Justice, C., and Liu, H. 1994. Development of vegetation and soil indices for MODIS-EOS. Remote Sensing of Environment, 49, pp.224–234.

Huete, A.R., Justice, C., and van Leeuwen, W. 1999. MODIS Vegetation Index (MOD13) Algorithm Theoretical Basis Document, NASA Goddard Space Flight Center, http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf, pp.120.

Huete, A.R., Liu, H.Q., Batchily, K., and van Leeuwen, W.J.D. 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. Remote Sens. Environ., 59, pp.440–451.

IPCC. 2003. Intergovernmental Panel on Climate Change. Good Practice Guidance on Land Use, Land-Use Change and Forestry. Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T. and Tanabe, K. (eds.). National Greenhouse Gas Inventories Programme. Japan: Institute for Global Environmental Strategies (IGES).

IPCC. 2013. Summary for policymakers. In, Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Ivits, E., Lamb, A., Langar, F., Hemphill, S., Koch, B. 2008. Orthogonal transformation of segmented SPOT5 images. Photogramm. Eng. Remote. Sens., 74, pp.1351–1364.

Jarvis, P., and Linder, S. 2000. Botany: constraints to growth of boreal forests. Nature, 405(6789), pp.904–905.

Jensen, J.R. 2007. Remote Sensing of the Environment: An Earth Resource Perspective; second edition. New Jersey: Pearson Prentice Hall.

Jin, H., and Eklundh, L. 2014. A physically based vegetation index for improved monitoring of plant phenology. Remote Sensing of Environment, 152, pp.512–525.

Jingyong, Z., Wenjie, D., Congbin, F., and Lingyun, W. 2003. The influence of vegetation cover on summer precipitation in China: a statistical analysis of NDVI and climate data. Advances in Atmospheric Sciences, 20, pp.1002–1006.

Kauth, R.J., and Thomas, G.S. 1976. The tasseled cap–a graphic description of the spectral-temporal development of agricultural crops as seen in Landsat. In, Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, Indiana.

Kauth, R.J., Lambeck, P.F., Richardson, W.R., Thomas, G.S., and Pentland, A.P. 1979. Feature extraction applied to agricultural crops as seen by Landsat. In, Proc. Technical Sessions, The LACIE Symposium, JSC 16015, NASA Johnson Space Center, Houston, Texas, vol ll., pp.705–721.

Kawabata, A., Ichii, K., and Yamaguchi, Y. 2001. Global monitoring of interannual changes in vegetation activities using NDVI and its relationships to temperature and precipitation. International Journal of Remote Sensing, 22, pp.1377–1382.

Keith, H., Mackey, B., Berry, S., Lindenmayer, D., and Gibbons, P. 2010. Estimating carbon carrying capacity in natural forest ecosystems across heterogeneous landscapes: addressing sources of error. Glob. Chang. Biol., 16, pp.2971–2989.

Kembel, S.W., and Dale, M.R.T. 2006. Within-stand spatial structure and relation of boreal canopy and understory vegetation. J. Veg. Sci., 17, pp.783–790.

Kennedy, R.E., Braaten, J, Yang, Z., Nelson, P., and Duane, M. 2013. LandTrendr Users Guide V3.0.

Kennedy, R.E., Cohen, W.B., and Schroeder, T.A. 2007. Trajectory-based change detection for automated characterization of forest disturbance dynamics. Remote Sens. Environ., 110, pp.370–386.

Kennedy, R.E., Yang, Z., and Cohen, W.B. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. Remote Sens. Environ., 114, pp.2897–2910.

Kennedy, R.E., Yang, Z., Cohen, W.B., Pfaff, E., Braaten, J., and Nelson, P. 2012. Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan. Remote Sens. Environ., 122, pp.117–133.

Kennedy, R.E., Yang, Z., Gorelick, N., Braaten, J., Cavalcante, L., Cohen, W.B., and Healey, S. 2018. Implementation of the LandTrendr Algorithm on Google Earth Engine. Remote Sensing., 10, pp.691.

Key, C.H., and Benson, N.C. 2005. Landscape assessment: sampling and analysis methods. In, D.C. Lutes, R.E. Keane, and J.F. Caratti (Eds), FIREMON: Fire Effects Monitoring and Inventory System. General Technical Report, RMRS-GTR-164. USDA Forest Service, Rocky Mountain Research Station, Ogden, Utah.

Keys, L.D., Schmidt, N.J., and Phillips, B.E. 1990. A prototype example of sensor fusion used for a siting analysis. Technical Papers 1990, ACSM-ASPRS Annual Convention, Image Processing and Remote Sensing, 4, pp.238–249.

Kurz, W.A., Stinson, G., and Rampley, G. 2008b. Could increased boreal forest ecosystem productivity offset carbon losses from increased disturbances? Philosophical Transactions of the Royal Society of London B: Biological Sciences, 363, pp.2259–2268.

Kushida, K., Kim, Y., Tsuyuzaki, S., and Fukuda, M. 2009. Spectral vegetation indices for estimating shrub cover, green phytomass and leaf turnover in a sedge-shrub tundra. International Journal of Remote Sensing, 30, pp.1651–1658.

Lambert, J., Drenou, C., Denux, J.P., Balent, G., and Cheret, V. 2013. Monitoring forest decline through remote sensing time series analysis. GISci. Remote Sens., 50, pp.437–457.

LandTrendr Users Guide V3.0. 2013. http://landtrendr.forestry.oregonstate.edu.

Lanir, J. 2007. Comparing multispectral image fusion methods for a target detection task. Optical Engineering, 46, pp.66402-1.

Lemprière, T.C., Bernier, P.Y., Carroll, A.L., Flannigan, M.D., Gilsenan, R.P., McKenney, D.W., Hogg, E.H., Pedlar, J.H., and Blain, D. 2008. The importance of forest sector adaptation to climate change (Vol. 416). Northern Forestry Centre.

Lenton, T.M. 2013. Environmental tipping points. Annual Review of Environment and Resources, 38, pp.1–29.

Lillesand, T.M., Kiefer, R.W., and Chipman, J.W. 2008. Remote sensing and image interpretation. (6th ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Liu, M. L. 2001: Research of land use/land cover change and terrestrial ecosystem vegetation carbon and production cover China (in Chinese). Ph.D. dissertation, Institute of Remote Sensing Applications, Chinese Academy of Sciences.

Liu, Y., Xiao, J., Ju, W., Xu, K., Zhou, Y., and Zhao, Y. 2016. Recent trends in vegetation greenness in China significantly altered annual evapotranspiration and water yield. Environ. Res. Lett., 11, pp.094010.

Lopez Garcia, M., and Caselles, V. 1991. Mapping burns and natural reforestation using Thematic Mapper data. Geocarto International 6, pp.31–37.

Lu, D., Mausel, P., Brondizio, E., and Moran, E. 2004. Change detection techniques. International Journal of Remote Sensing, 25, pp.2365−2407.

Lu, L., Li, X., Veroustraete, F., Kang, E., and Wang, J. 2009. Analysing the forcing mechanisms for net primary productivity changes in the Heihe River Basin, northwest China. Int. J. Remote Sens., 30, pp.793–816.

Luckman, B.H., Youngblut, D.K., Wilson R.J.S., Watson, E., Colenutt, M.E., and St. George, R.S. 2004. Dendroclimatology in the Canadian Cordillera. In, Meeting on Tree Rings and Climate: Sharpening the Focus. Marriott University Park, Tucson, AZ.

Luigi-Dragotti, P., Poggi, G., and Ragozini, A. R. P. 2000. Compression of multispectral images by three-dimensional SPIHT algorithm. IEEE Trans. Geosci. Remote Sensing, 38, pp. 416–428.

Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J.G., and Worthy, L.D. 2006. Land-cover change detection using multi-temporal MODIS NDVI data. Remote Sens. Environ., 105, pp.142–154.

Lyon, J.G., Yuan, D., Lunetta, R.S., and Elvidge, C.D. 1998. A change detection experiment using vegetation indices. Photogrammetric Engineering and Remote Sensing, 64(2), pp.143−150.

Margono, B.A., Turubanova, S., Zhuravleva, I., Potapov, P., Tyukavina, A., Baccini, A., Goetz, S., Hansen, M.C. 2012. Mapping and monitoring deforestation and forest degradation in Sumatra (Indonesia) using Landsat time series data sets from 1990 to 2010. Environ. Res. Lett., 7, pp.034010.

Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., and Lim, T.-K. 2006. A Landsat surface reflectance dataset for North America, 1990–2000. IEEE Geosci. Remote Sens. Lett., 3, pp.68–72.

Maselli, F., Romanelli, S., Bottai, L., and Zipoli, G. 2003. Use of NOAA-AVHRR NDVI images for the estimation of dynamic fire risk in Mediterranean areas. Remote Sensing of Environment, 86, pp.187−197.

Matsushita, B., M. Xu, J. Chen, S. Kameyama, and M. Tamurd, 2004. Estimation of regional net primary productivity (NPP) using a process-based ecosystem model: How important is the accuracy of climate data? Ecol. Modell., 178, pp. 371–388.

Melton, J., et al. 2013. Present state of global wetland extent and wetland methane modelling: conclusions from a model intercomparison project (WETCHIMP). Biogeosciences, 10, pp.753–788.

Meneses, B.M. 2021. Vegetation recovery patterns in burned areas assessed with Landsat 8 OLI imagery and environmental biophysical data. Fire, 4, pp.76.

Meng, D., Li, X., Zhao, W., and Gong, H. 2009. Quantitative exploration of the mechanisms behind the urban thermal environment in Beijing. Progress in Natural Science, 19, pp.1757–1763.

Meng, J., Du, X., and Wu, B. 2013. Generation of high spatial and temporal resolution NDVI and its application in crop biomass estimation. International Journal of Digital Earth, 6, pp.203–218.

Mikwa, J.F. Gossens, R. and Defourny, P. 2016. Forest degradation, a methodological approach using remote sensing techniques: A review. International Journal of Innovation and Scientific Research, 24(1), pp.161–178.

Miller, J.D., and Yool, S.R. 2002. Mapping forest post-fire canopy consumption in several overstory types using multi-temporal Landsat TM and ETM data. Remote Sensing of Environment, 82(2–3), pp. 481−496.

MODIS GPP/NPP Products. https://modis.gsfc.nasa.gov/data/dataprod/mod17.php

Monteith, J.L., 1972. Solar radiation and productivity in tropical ecosystems. J. Appl. Ecol., 9(3), pp.747–766.

Mouat, D.A., Mahin, G.G., and Lancaster, J. 1993. Remote sensing techniques in the analysis of change detection. Geocarto International, 2, pp.39–50.

Nabuurs, G.J., Masera, O., Andrasko, K., Benitez-Ponce, P., Boer, R., Dutschke, M., Elsiddig, E., Ford-Robertson, J., Frumhoff, P., Karjalainen, T., Krankina, O., Kurz, W.A., Matsumoto, M., Oyhantcabal, W., Ravindranath, N.H., Sanz Sanchez, M.J., and Zhang, X. 2007. Forestry. In, B. Metz, O.R. Davidson, P.R. Bosch, R. Dave, L.A. Meyer (Eds.), Climate Change 2007: Mitigation. Contribution of Working Group

III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

NASA. MODIS gap-filled yearly NPP products (MOD17A3HGF-V061). modis.gsfc.nasa.gov.

Natural Resources Canada. 2016. The state of Canada's forests: Annual report 2015. Government of Canada, Ottawa, Ontario, Canada.

Nelson, P., Pfaff, E., and Kennedy, R. 2010. Standard operating procedure #2: Preprocessing image stacks. In, R. Kennedy, A. Kirschbaum, U. Gafvert, P. Nelson, Z. Yang, W. Cohen, E. Pfaff, and B. Gholson (Eds.), Landsat-based monitoring of landscape dynamics in the national parks of the Great Lakes Inventory and Monitoring Network (Version 1.0). Natural Resource Report NPS/GLKN/NRR—2010/221. Fort Collins, Colorado: National Park Service.

Nemani, R.R., Keeling, C.D., Hashimoto, H., Jolly, W.M., Piper, S.C., Tucker, C.J., Myneni, R.B., and Running, S.W. 2003. Climate-driven increases in global terrestrial net primary production from 1982 to 1999. Science, 300, pp.1560–1563.

Netherer, S., and Schopf, A. 2010. Potential effects of climate change on insect herbivores in European forests— General aspects and the pine processionary moth as specific example. Forest Ecology and Management, 259, pp.831–838.

Neuvonen, S., and Virtanen, T. 2015. Abiotic factors, climate variability and forest insect pests. In, C. Björkman and P. Niemelä (Eds.), Climate Change and Insect Pests (pp. 235–247). Boston: CABI.

Nunez, J., Otazu, X., Fors, O., Prades, A., Pala, V., and Arbiol, R. 1999. Multiresolution-based image fusion with additive wavelet decomposition. IEEE Trans. Geosci. Remote Sensing, 37, pp.1204–1211.

Pan, Y., et al. 2013. The structure, distribution, and biomass of the world's forests. Annu. Rev. Ecol. Evol. Syst., 44, pp.593–622.

Paruelo, J.M., et al. 1997 ANPP estimates from NDVI for the central grassland region of the United States. Ecology, 78(3), pp.953–958.

Peng, D. L., Huang, J.-F., Cai, C.-X., Deng, R., and Xu, J.-F. 2008. Assessing the response of seasonal variation of net primary productivity to climate using remote sensing data and geographic information system techniques in Xinjiang. J. Integr. Plant Biol., 50, pp.1580–1588.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J., and Stenseth, N.C. 2005. Using the satellite-derived NDVI to assess ecological responses to environmental change. Trends in Ecology & Evolution, 20, pp.503–510.

Pflugmacher, D., Cohen, W.B., Kennedy, R.E. 2012. Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. Remote Sens. Environ., 122, pp.146–165.

Pflugmacher, D., Cohen, W.B., Kennedy, R.E., Yang, Z. 2013. Using Landsat-derived disturbance and recovery history and lidar to map forest biomass dynamics. Remote Sens. Environ., 151, pp.124–137.

Pinto, C.T., Haque, M.O., Micijevic, E., Helder, D.L. 2019. Landsats 1–5 multispectral scanner system sensors radiometric calibration update. IEEE Trans. Geosci. Remote Sens., 57(10), pp.7378–7394.

Pohl, C. 1996. Geometric aspects of multisensor image fusion for topographic map updating in the humid tropics. ITC publication No. 39, ISBN 90 6164 121 7, 160 pages, 2 image maps.

Pohl, C., and Van Genderen, J. L. 1998. Review article multisensor image fusion in remote sensing: concepts, methods and applications. International Journal of Remote Sensing, 19, pp.823–854.

Pohl, C., and Van Genderen, J.L. 1993. Geometric integration of multi-image information. Space at the Service of our Environment, Proceedings of the Second ERS-1 Symposium, 11–14 October 1993, Hamburg, Germany, ESA SP-361 (Paris: European Space Agency), pp. 1255–1260.

Potter, C. S., Klooster, S.A., and Myneni, R. 2003. Continental scale comparisons of terrestrial carbon sinks estimated from satellite data and ecosystem modeling 1982–1998. Global Planet Change, 393, 201–213.

Prentice, I.C., et al. 2001. The carbon cycle and atmospheric carbon dioxide. Cambridge University Press.

Price, D.T., Alfaro, R.I., Brown, K.J., Flannigan, M.D., Fleming, R.A., Hogg, E.H., Girardin, M. P., Lakusta, T., Johnston, M., McKenney, D.W., Pedlar, J., Stratton, T., Sturrock, R., Thompson, I., Trofymow, J.A., and Venier, L.A. 2013. Anticipating the consequences of climate change for Canada's boreal forest ecosystems 1. Environmental Reviews, 21(4), pp.322–365.

Raich, J., Rastetter, E., Mellilo, J., Kicklighter, D., Steudler, P., Peterson, B., Grace, A.L., Moore, B. III, and Vorosmarty, C.J. 1991. Potential net primary productivity in South America: application of a global model. Ecological Applications, 1, pp.399–429.

Ramon Solano, R., Didan, K., Jacobson, A., and Huete, A. 2010. Modis Vegetation Index User's Guide. Tucson: The University of Arizona.

Rautiainen, M., Heiskanen, J., Eklundh, L., Mottus, M., Lukeš, P., and Stenberg, P. 2010. Ecological applications of physically based remote sensing methods. Scandinavian Journal of Forest Research, 25, pp.325–339.

Raynolds, M., Walker, D., and Maier, H. 2006a. NDVI patterns and phytomass distribution in the circumpolar Arctic. Remote Sensing of Environment, 102, pp.271–281.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., and Ohlen, D.O. 1994. Measuring phenological variability from satellite imagery. Journal of Vegetation Science, 5, pp.703–714.

Richards, J.A. 1999. Remote Sensing Digital Image Analysis: An Introduction. Berlin, Germany: Springer-Verlag.

Roerink, G., Menenti, M., Soepboer, W., and Su, Z. 2003. Assessment of climate impact on vegetation dynamics by using remote sensing. Physics and Chemistry of the Earth, Parts A/B/C, 28, pp.103–109.

Rogan, J., and Chen, D. 2004. Remote sensing technology for mapping and monitoring land-cover and land-use change. Prog. Plan., 61, pp.301–325.

Rogers, R.H., and Wood, L. 1990. The history and status of merging multiple sensor data: an overview. Technical Papers 1990, ACSM-ASPRS Annual Convention, Image Processing and Remote Sensing, 4, pp.352–360.

Romero-Sanchez, M.E., and Ponce-Hernandez, R. 2015. The assessment of forest degradation in dry forested lands: Mapping regional trend indicators of degradation in the Yucatan Peninsula, Mexico with satellite data. In, Hubert, B., Broin, M. (Eds.), 3rd UNCCD Scientific Conference: "Combating Desertification/land Degradation and Drought for Poverty Reduction and Sustainable Development: The Contribution of Science, Technology, Traditional Knowledge and Practices." Agropolis International, Montpelier, France, pp. 284–285.

Romero-Sanchez, M.E., and Ponce-Hernandez, R. 2017. Assessing and monitoring forest degradation in a deciduous tropical forest in Mexico via remote sensing indicators. Forests, 8, pp.302.

Rouse, J.W., Haas, R.H., Schell, J.A., and Deering, D.W. 1974. Monitoring vegetation systems in the Great Plains with ERTS. NASA Spec. Publ., 351, pp.309.

Roy, D.P., Wulder, M.A., Loveland, T.R., Carabajal C.C., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., et al. 2014. Landsat-8: science and product vision for terrestrial global change research. Remote Sensing of Environment, 145, pp.154–172.

Running, S.W., and Zhao, M.S. 2015. User's Guide. Daily GPP and Annual NPP (MOD17A2/A3) Products NASA Earth Observing System MODIS Land Algorithm. Version 3.0 for Collection 6. https://lpdaac.usgs.gov/sites/default/ files/public/product_documentation/mod17_user_guide.pdf.

Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M., Reeves, M., and Hashimoto, H. 2004. A continuous satellite-derived measure of global terrestrial primary production. Bioscience, 54(6), pp.547–560.

Running, S.W., Thornton, P.E., Nemani, R.R., and Glassy, J.M. 2000. Global terrestrial gross and net primary productivity from the Earth Observation system. In, O.E. Sala, R.B. Jackson, H.A. Mooney, and R.W. Howart (Eds.), Methods in Ecosystem Science. New York: Springer, pp. 44–57.

Saunders, S.C., Chen, J., Drummer, T.D., Crow, T.R., Brosofske, K.D., and Gustafson, E.J. 2002. The patch mosaic and ecological decomposition across spatial scales in a managed landscape of northern Wisconsin, USA. Basic Appl. Ecol., 3, pp.49–64.

Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., and Walker, B. 2001. Catastrophic shifts in ecosystems. Nature, 413, pp.591–596.

Scholes, R.J., and Hall, D. 1996. The carbon budget of tropical savannas, woodlands and grasslands. SCOPE—Scientific Committee on Problems of the Environment International Council of Scientific Unions, 56, pp.69–100

Schuur, E.A. 2003. Productivity and global climate revisited: the sensitivity of tropical forest growth to precipitation. Ecology, 84(5), pp.1165–1170.

Seaquist, J.W., Olsson, L., and Ardö, J. 2003. A remote sensing-based primary production model for grassland biomes. Ecol. Modell., 169, pp.131–155.

Seidl, R., Schelhaas, M.J., Rammer, W., and Verkerk, P.J. 2014. Increasing forest disturbances in Europe and their impact on carbon storage. Nature Climate Change, 4, pp.930–930.

Sellers, P., Berry, J., Collatz, G., Field, C., and Hall, F. 1992. Canopy reflectance, photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. Remote Sensing of Environment, 42, pp.187–216.

Shen, S.S. 1990. Summary of types of data fusion methods utilized in workshop papers. Multisource Data Integration in Remote Sensing, Proceedings of Workshop held in Maryland, U.S.A., 14–15 June 1990, NASA Conference Publication 3099, pp. 145–149.

Shettigara, V.K. 1992. A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set. Photogrammetric Engineering and Remote Sensing, 58, pp.561–567.

Sidhu, N., Pebesma, E., and Câmara, G. 2018. Using Google Earth Engine to detect land cover change: Singapore as a use case. Eur. J. Remote Sens., 51, pp.486–500.

Simhadri, K.K., Iyengar, S.S, Holyer, R.J., Lybanon, M., and Zachary Jr, J.M. 1998. Wavelet-based feature extraction from oceanographic images. IEEE Trans. Geosci. Remote Sensing, 36, pp. 767–778.

Singh, A. 1989. Digital change detection techniques using remotely-sensed data. International Journal of Remote Sensing, 10, 989–1003.

Singh, A., and Harrison, A. 1985. Standardized principal components. International Journal of Remote Sensing, 6, pp.883–896.

Singh, R.P., Roy, S., and Kogan, F. 2003. Vegetation and temperature condition indices from NOAA AVHRR data for drought monitoring over India. International journal of remote sensing, 24, pp.4393-4402.

Stocks, B.J., Fosberg, M.A., Lynham, T.J., Mearns, L., Wotton, B.M., Yang, Q., Jin, J.Z., Lawrence, K., Hartley, G.R., Mason, J.A., and McKenney, D.W. 1998. Climate change and forest fire potential in Russian and Canadian boreal forests. Climatic Change, 38(1), pp.1–13.

Stocks, B.J., Mason, J.A., Todd, J.B., Bosch, E.M., Wotton, B.M., Amiro, B.D., Flannigan, M.D., Hirsch, K.G., Logan, K.A., Martell, D.L., and Skinner, W.R. 2002. Large forest fires in Canada, 1959–1997. Journal of Geophysical Research: Atmospheres (1984–2012), 107(D1), FFR-5.

Susmitha V., and Pancham S. 2009. A novel architecture for wavelet based image fusion. World Academy of Science, Engineering and Technology, 57.

Tait, A., and Zheng, X. 2003. Mapping frost occurrence using satellite data. Journal of Applied Meteorology, 42, pp.193−203.

Tao, F.L., and Zhang, Z. 2010. Dynamic responses of terrestrial ecosystems structure and function to climate change in China. J. Geophys. Res., 115, pp.G03003.

Taylor-Sakyi, K. 2012. Big data: understanding big data. arXiv 2016, arXiv:1601.04602, preprint. 19.

Thoha, A.S. 1998. Assessment of forest fire danger by using Keetch–Byram drought index in Sumberkima Bali Province. Thesis. Bogor: Bogor Agricultural University.

Thompson, I.D., Guariguata, M.R., Okabe, K., Bahamondez, C., Nasi, R., Heymell, V., and Sabogal, C. 2013. An operational framework for defining and monitoring forest degradation. Ecol. Soc., 18, pp.18.

Tucker, C.J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment, 8(2), pp.127−150.

Tucker, C.J., and Sellers, P.J. 1986. Satellite remote-sensing of primary production. International Journal of Remote Sensing, 7(11), pp.1395−1416.

Upadhyay, T.P., Sankhayan, P.L., and Solberg, B. 2005. A review of carbon sequestration dynamics in the Himalayan region as a function of land-use change and forest/soil degradation with special reference to Nepal. Agric. Ecosyst. Environ., 105, pp.449–465.

USGS. 2015. Product Guide: Landsat Surface Reflectance-Derived Spectral Indices.

Vekkot, S., and Shukla, P. 2009. A novel architecture for wavelet based image fusion. World Academy of Science, Engineering and Technology, 3, pp.2205–2210.

Verbesselt, J., Hyndman, R., Newnham, G., and Culvenor, D. 2010. Detecting trend and seasonal changes in satellite image time series. Remote Sens. Environ., 114, pp.106–115.

Verbesselt, J., Zeileis, A., Herold, M. 2012. Near real-time disturbance detection using satellite image time series. Remote Sens. Environ., 123, pp.98–108.

Verstraeten, W.W., Veroustraete, F., and Feyen, J. 2006. On temperature and water limitation of net ecosystem productivity: implementation in the C-Fix model. Ecol. Modell., 199, pp.4–22.

Viana-Soto, A., Aguado, I., and Martínez, S. 2017. Assessment of post-fire vegetation recovery using fire severity and geographical data in the Mediterranean region (Spain). Environments, 4, pp.90.

Vierling, L.A., Deering, D.W., and Eck, T.F. 1997. Differences in Arctic tundra vegetation type and phenology as seen using bidirectional radiometry in the early growing season. Remote Sensing of Environment, 60, pp.71–82.

Walker, D.A., Epstein, H.E., Jia, G.J., Balser, A., Copass, C., Edwards, E.J., Gould, W.A., Hollingsworth, J., Knudson, J., Maier, H.A., Moody, A., and Raynolds, M.K. 2003. Phytomass, LAI, and NDVI in northern Alaska: relationships to summer warmth, soil pH, plant functional types, and extrapolation to the circumpolar Arctic. Journal of Geophysical Research-Atmospheres, 108.

Wang, Q., Watanabe, M., Hayashi, S., and Murakami, S. 2003b. Using NOAA AVHRR data to assess flood damage in China. Environmental Monitoring and Assessment, 82, pp.119–148.

Wang, Q., Zhao, P., Ren, H., and Kakubari, Y. 2008. Spatiotemporal dynamics of forest net primary production in China over the past two decades. Global Planet Change, 61, pp.267–274.

Weng, Q., Fu, P., and Gao, F. 2014. Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. Remote Sensing of Environment, 145, pp.55–67.

Werness, S.A., Wei, S.C., and Carpinella, R. 1994. Experiments with wavelets for compression of SAR data. IEEE Trans Geosci. Remote Sensing, 32, pp.197–201.

Weydahl, D.J. 1992. Temporal change detection in ERS-1 SAR images. International Space Year: Space Remote Sensing, Proceedings 12th Annual IEEE International Geoscience and Remote Sensing Symposium (IGARSS '92), Houston, U.S.A., edited by R. Williamson with collaboration of T. Stein (New York), pp. 1346–1348.

Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S.N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P.S., Vermote, E.F., Vogelmann, J., Wulder, M.A., and Wynne, R. 2008. Free access to Landsat imagery. Science, 320, pp.1011.

Woodward, F., Lomas, M., and Kelly, C. 2004. Global climate and the distribution of plant biomes. Philos. Trans. R. Soc. B. Biol. Sci., 359(1450), pp.1465–1476.

Wulder, M. 2008b. Landsat continuity: issues and opportunities for land cover monitoring. Remote Sensing of Environment, 112, pp.955–969.

Wulder, M.A., Butson, C.R., and White, J.C. 2008. Cross-sensor change detection over a forested landscape: options to enable continuity of medium spatial resolution measures. Remote Sensing of Environment, 112, pp.796−809.

Wulder, M.A., Loveland, T.R., Roy, D.P., Crawford, C.J., et al. 2019. Current status of Landsat program, science, and applications. Remote Sens. Environ., 225, pp.127–147.

Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., and Woodcock, C.E. 2012. Opening the archive: how free data has enabled the science and monitoring promise of Landsat. Remote Sens. Environ., 122, pp.2–10.

Xia, J., et al. 2014. Spatio-temporal patterns and climate variables controlling of biomass carbon stock of global grassland ecosystems from 1982 to 2006. Remote Sens., 6(3), pp.1783–1802.

Xu, L., Myneni, R.B., Chapin III, F.S., Callaghan, T.V., Pinzon, J.E., Tucker, C.J., Zhu, Z., Bi, J., Ciais, P., Tommervik, H., Euskirchen, E.S., Forbes, B.C., Piao, S.L., Anderson, B.T., Ganguly, S., Nemani, R.R., Goet, S.J., Beck, P.S.A., Bunn, A.G., Cao, C., and Stroeve, J.C. 2013. Temperature and vegetation seasonality diminishment over northern lands. Nature Climate Change Letters, 1836, pp.1–6.

Yang, Y., Erskine, P.D., Lechner, A.M., Mulligan, D., Zhang, S., and Wang, Z. 2018. Detecting the dynamics of vegetation disturbance and recovery in surface mining area via Landsat imagery and LandTrendr algorithm. J. Clean. Prod., 178, pp.353–362.

Yuan, W., Liu, S., Zhou, G., Zhou, G., Tieszen, L.L., Baldocchi, D., Bernhofer, C., Gholz, H., Goldstein, A.H., Goulden, M.L., Hollinger, D.Y., Hu, Y., Law, B.E., Stoy, P.C., Vesala, T., and Wofsy, S.C. 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. Agric. For. Meteorol., 143, pp.189–207.

Zanchetta, A., Bitelli, G., Karnieli, A. 2016. Monitoring desertification by remote sensing using the tasselled cap transform for long-term change detection. Natural Hazards, 83(1), pp.115.

Zhang, J. 2010. Multi-source remote sensing data fusion: status and trends. International Journal of Image and Data Fusion, 1, pp.5–24.

Zhang, W., Li, A., Jin, H., Bian, J., Zhang, Z., Lei, G., Qin, Z., and Huang, C. 2013. An enhanced spatial and temporal data fusion model for fusing Landsat and MODIS surface reflectance to generate high temporal Landsat-Like Data. Remote Sensing, 5, pp.5346–5368.

Zhang, Y. 2004. Understanding image fusion. Photogramm. Eng. Remote Sens., 70, pp. 657–661.

Zhang, Y. 2008. Methods for image fusion quality assessment – a review, comparison and analysis. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVII, pp.1101–1109.

Zhang, Z., and Blum, R.S. 1995. Region-based image fusion scheme for concealed weapon detection. In Proceedings of the 31st Annual Conference on Information Sciences and Systems.

Zheng, L., Chan, A.K., Liu, S., Smith, W., and Holyer, R.J. 1999. Directional clutter removal of aerial digital images using X-ray wavelet transform and Markov random field. IEEE Trans. Geosci. Remote Sensing, 37, pp. 2181–2191.

Zhou, L., Kaufmann, R., Tian, Y., Myneni, R., and Tucker, C. 2003. Relation between interannual variations in satellite measures of northern forest greenness and climate between 1982 and 1999. Journal of Geophysical Research: Atmospheres, 108(D1), pp.4004.

Zhou, Y., and Hecht-Nielsen, R. 1993. Target recognition using multiple sensors. Proc. IEEE Workshop: Neural Networks for signal Processing, pp. 411–420.

Zhu, L., and Southworth, J. 2013. Disentangling the relationships between net primary production and precipitation in southern Africa savannas using satellite observations from 1982 to 2010. Remote Sens., 5(8), pp.3803–3825.

Zhu, Z. 2017. Change detection using Landsat time series: a review of frequencies, preprocessing, algorithms, and applications. ISPRS J. Photogramm. Remote Sens., 130, pp.370–384.

Zhu, Z., and Woodcock, C.E. 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens. Environ., 118, pp.83–94.

Appendix 3

3A: File name: run_ledaps_landtrendr_processor_1728_nbr_eval

```
retall
;this batch file controls all of the ledaps landtrendr preprocessing steps
;---please state the landsat path/row ID and path to the data as described following each variable---
ppprrr = '017028'
path = 'C:\017028\'
useareafile = 'C:\017028\images\studyarea.bsq'
segparamstxt = 'C:\myworkspace\myproject\paramfiles\nbr_segmentation_parameters.txt'
template_hdr = 'C:\myworkspace\myproject\paramfiles\utm_template_headerfile_LT5.hdr'
label_parameters_txt = 'C:\myworkspace\myproject\paramfiles\eval_label_params.txt'
class_code_txt = 'C:\myworkspace\myproject\paramfiles\eval_class_codes.txt'
;---- if running ftv with post-processing detrending, need to specfy the
; file with the percent cover, etc. values
; otherwise specify 'none'
;post_process_file = '/projectnb/trenders/helperfiles/nbr_label_parameters.txt'
;PROCESSING SWITCHES
;1=do this, 0=don't do this, or as described
resume_segmentation  = 0  ;if segmentation crashed mid-process set this to 1 to begin on the chunk it let off, else leave at 0
segmentation_eval    = 1  ;runs landtrendr in evaluation mode
segmentation         = 0  ;creates segmentation outputs
fit_to_vertices      = 0  ;use 1 to run bgw, 2 to run b5,b4,b3, 3 to run both, 5 for NBR and BGW-Hoque
dist_rec_snapshots   = 0  ;create disturbance and recovery slice outputs
dark_seg_outputs     = 0  ;creates an output used to make a forest-nonforest mask
progressbaryesno     = 1  ;set to 1 to have a graphical progress bar during segmentation, 0 to use text

;(OPTIONAL) ADJUST THE IMAGES TO RUN THROUGH SEGMENTATION
image_list = [0] ;list image dates to leave in or take out of segmentation - leave as 0 if you want to use all images
  ;image date format: single: [1984226] multiple: [1984226,1985196,1992216]
image_list_type = 1 ;1 = exclude above dates ::::::: 2 = run only on above dates
run_params = {ppprrr:ppprrr,$
  path:path,$
  segmentation_eval:segmentation_eval,$
  segparamstxt:segparamstxt,$
  segmentation:segmentation,$
  resume:resume_segmentation,$
  useareafile:useareafile,$
```

```
  fit_to_vertices:fit_to_vertices,$
  dist_rec_snapshots:dist_rec_snapshots,$
  dark_seg_outputs:dark_seg_outputs,$
  image_list:image_list,$
  image_list_type:image_list_type,$
  template_hdr:template_hdr, $
  progressbaryesno:progressbaryesno, $
  label_parameters_txt:label_parameters_txt, $
  class_code_txt:class_code_txt}
  .run tbcd_v2
ledaps_landtrendr_processor, run_params
```

3B: File: run_ledaps_landtrendr_processor_1828_wetness [MSS image pathe=018, row=028**]**

```
retall
;this batch file controls all of the ledaps landtrendr preprocessing steps
;---please state the landsat path/row ID and path to the data as described following each variable---
ppprrr = '018028'
path = 'C:\018028\'
useareafile = 'C:\018028\images\useareamss.bsq'
segparamstxt = 'C:\myworkspace\myproject\paramfiles\wetness_segmentation_parameters.txt'
template_hdr = 'C:\myworkspace\myproject\paramfiles\utm_template_headerfile_MSS.hdr'
;---- if running ftv with post-processing detrending, need to specfy the
; file with the percent cover, etc. values
; otherwise specify 'none'
;post_process_file = 'C:\myworkspace\myproject\paramfiles\wetness_label_parameters.txt'
;PROCESSING SWITCHES
;1=do this, 0=don't do this, or as described
resume_segmentation  = 0  ;if segmentation crashed mid-process set this to 1 to begin on the chunk it let off, else
leave at 0
segmentation_eval    = 0  ;runs landtrendr in evaluation mode
segmentation         = 1  ;creates segmentation outputs
fit_to_vertices      = 1  ;use 1 to run bgw, 2 to run b5,b4,b3, 3 to run both
dist_rec_snapshots   = 0  ;create disturbance and recovery slice outputs
dark_seg_outputs     = 1  ;creates an output used to make a forest-nonforest mask
progressbaryesno     = 0  ;set to 1 to have a graphical progress bar during segmentation, 0 to use text

;(OPTIONAL) ADJUST THE IMAGES TO RUN THROUGH SEGMENTATION
image_list = [0] ;list image dates to leave in or take out of segmentation - leave as 0 if you want to use all images
  ;image date format: single: [1984226] multiple: [1984226,1985196,1992216]
image_list_type = 1 ;1 = exclude above dates ::::::: 2 = run only on above dates

run_params = {ppprrr:ppprrr,$
```

```
  path:path,$
  segmentation_eval:segmentation_eval,$
  segparamstxt:segparamstxt,$
  segmentation:segmentation,$
  resume:resume_segmentation,$
  useareafile:useareafile,$
  fit_to_vertices:fit_to_vertices,$
  dist_rec_snapshots:dist_rec_snapshots,$
  dark_seg_outputs:dark_seg_outputs,$
  image_list:image_list,$
  image_list_type:image_list_type,$
  template_hdr:template_hdr, $
  progressbaryesno:progressbaryesno}
  .run tbcd_v2
ledaps_landtrendr_processor, run_params
```

3C: File: ndvi_segmentation_parameters

```
run_name            = paramset01
base_index          = ndvi
background_val      = 0
divisor             = 1
minneeded           = 3
kernelsize          = 3
pval                = 0.05
fix_doy_effect      = 1
max_segments        = 5
recovery_threshold  = 1
skipfactor          = 1
desawtooth_val      = 0.9
distweightfactor    = 1
vertexcountovershoot = 3
bestmodelproportion  = 0.75
mask_image          = na
ulx                 = na
uly                 = na
lrx                 = na
lry                 = na
```

Note: Applied/replaced other FDIs instead of NDVI such as base_index = nbr, wetness, etc.

3D: File: nbr_label_parameters

```
static_model        = static_nbr_model_pilot_scenes
change_model        = none
pct_tree_loss1      = 10
```

```
pct_tree_loss20      = 3
pre_dist_cover       = 20
pct_tree_gain        = 5
collapse_dist_angle  = 15
collapse_recv_angle  = 15
run_name             = nbr_lt_labels
merge_recovery       = yes
extract_tc_ftv       = no
use_relative_mag     = yes
end_year             = -1
start_year           = -1
```

Note: Applied/replaced other FDIs instead of NBR such as run_name = ndvi_lt_labels, wetness_lt_labels, etc.

3E: File: run_lt_labelfilt_1728_wetness

```
;landtrendr post-segmentation change labeling\map creation and spatial filter and patch aggregation
;-------------inputs-------------
;full path to the diag.sav files in the outputs folder that you want to run labeling and filter on
diag_files=["C:\017028\outputs\wetness\LT_v2.1_wetness_017028_paramset01_20210314_004645_diag.sav"]
;full path to label parameter .txt files.  these correspond to the above daig_files so one must exist
;for for each, they do not have to be the same, but can be
label_parameters_txt=["C:\myworkspace\myproject\paramfiles\wetness_label_parameters.txt"]
;full path to a single class code file that defines what map outputs will be created
class_code_txt = "C:\myworkspace\myproject\paramfiles\wetness_label_codes.txt"
;full path to a projection template file
templatehdr = "C:\myworkspace\myproject\paramfiles\utm_template_headerfile_LT5.hdr"
;-------------run the program-------------
run_lt_label_and_filtering, diag_files, label_parameters_txt, class_code_txt, templatehdr
```

Note: Applied/replaced other FDIs instead of wetness.

3F: File: utm_template_headerfile_LT5

```
ENVI
description = {
  File Imported into ENVI.}
samples = 4411
lines   = 4279
bands   = 1
header offset = 0
file type = ENVI Standard
data type = 2
interleave = bsq
sensor type = Unknown
```

byte order = 0
x start = 2014
y start = 2739
map info = {UTM, 1.5000, 1.5000, 240375.000, 5128875.000, 3.0000000000e+001, 3.0000000000e+001, 18, North, WGS-84, units=Meters}
coordinate system string = {PROJCS["WGS_1984_UTM_Zone_18N",GEOGCS["GCS_WGS_1984",DATUM["D_WGS_1984",SPHEROID["WGS_1984",6378137.0,298.257223563]],PRIMEM["Greenwich",0.0],UNIT["Degree",0.0174532925199433]],PROJECTION["Transverse_Mercator"],PARAMETER["False_Easting",500000.0],PARAMETER["False_Northing",0.0],PARAMETER["Central_Meridian",-75.0],PARAMETER["Scale_Factor",0.9996],PARAMETER["Latitude_Of_Origin",0.0],UNIT["Meter",1.0]]}
wavelength units = Unknown

3G: File: studyarea

ENVI
description = {
Gain and Offset Correction [Tue Jan 19 23:07:17 2021]}
samples = 4411
lines = 4279
bands = 1
header offset = 0
file type = ENVI Standard
data type = 2
interleave = bsq
sensor type = Unknown
byte order = 0
x start = 2014
y start = 2739
map info = {UTM, 1.5000, 1.5000, 240375.000, 5128875.000, 3.0000000000e+001, 3.0000000000e+001, 18, North, WGS-84, units=Meters}

coordinate system string = {PROJCS["UTM_Zone_18N",GEOGCS["GCS_WGS_1984",DATUM["D_WGS_1984",SPHEROID["WGS_1984",6378137.0,298.257223563]],PRIMEM["Greenwich",0.0],UNIT["Degree",0.0174532925199433]],PROJECTION["Transverse_Mercator"],PARAMETER["False_Easting",500000.0],PARAMETER["False_Northing",0.0],PARAMETER["Central_Meridian",-75.0],PARAMETER["Scale_Factor",0.9996],PARAMETER["Latitude_Of_Origin",0.0],UNIT["Meter",1.0]]}

wavelength units = Unknown
data ignore value = 0
band names = {
Gainoff (Mask (Resize (Band 1:Exported_Reclass_b1_1984.tif):Resize_b1_1984):Studyarea_mask_b1_1984)}

CHAPTER FOUR

RELATIONSHIP BETWEEN CLIMATE VARIABLES AND FOREST DEGRADATION

Abstract

While there are relationships between traditional forest degradation indicators (FDIs), calculated from multispectral satellite imagery stacks, representing time-series, and climate variables, they have not been fully explored. In this study, therefore, FDIs were analyzed as indicators of the variations that occurred in the vegetation in the study area throughout time and space. A geostatistical analysis of climate was also conducted. To ascertain how historic climate variables related to FDIs, the computed FDIs from available multispectral image stacks were statistically related to historical climate variables, and in order to use such relationships to project FDIs, a multiple linear regression (MLR) model was developed. This research examined both the usefulness of some FDIs in explaining patterns of forest disturbance and established quantitative relationships between useful FDIs and climate variables.

In this study, forest degradation was monitored from 1972–2020. The composite forest degradation indicator (CFDI) was computed using principal component analysis (PCA) resulting from the integration of seven FDIs—normalized difference vegetation index (NDVI), net primary productivity (NPP), normalized burn ratio (NBR), enhanced vegetation index (EVI), tasseled cap brightness (TCB), tasseled cap greenness (TCG), and tasseled cap wetness (TCW), which were calculated from the satellite images (Landsat and MODIS) using ENVI. The relationships between remote sensing indicators (CFDI, NDVI, and NPP) and climate variables—temperature (T), precipitation (P), evapotranspiration (ETo), and moisture availability (estimated from two regimes: MA_1 and MA_2)—led to the creation of spatially-explicit maps of relevant climatic variables as well as long-term historical forest degradation maps, which were developed via the trajectory analysis computed with the LandTrendr algorithm (Kennedy et al., 2010).

This study focused on an analysis of the forest dynamic changes in the study area during the period from 1972–2020. The impacts of changes in T, P, ETo, MA_1, and MA_2 on forest vegetation conditions were studied from the perspective of the effects of temperature and water deficit. To understand the nature of the quantitative relationships between climate variables and FDIs, simple and MLR analysis was employed and an MLR model was developed to predict future dynamic changes in forest cover in the study area.

The findings for 1972–2020 indicated climate variables P, MA_1, and MA_2 had the strongest correlation (-0.95, -0.91, and -0.93, respectively) with CFDI, which enabled for computing an MLR model with a high coefficient of determination, R^2 (0.93), and low RMSE (0.28). Additionally, findings for 2001–2020 showed climate variables P, MA_1, and MA_2 had the most significant correlation with CFDI, NDVI, and NPP. This strong relationship allowed for the establishment of future models with a high coefficient of determination, R^2 (0.83, 0.80, 0.92, respectively), and low RMSE (0.39, 0.38, 0.26, respectively).

Finally, projected MLR models were established based on the highly correlated FDIs (CFDI, NDVI, and NPP) and climate variables (P, MA_1, and MA_2).

Trends in temperatures within and near the study area are showing a rise in daily average temperature, which was 4.9^0C, 5.1^0C, and 5.2^0C during the periods from 1971–2000, 1981–2010, and 2011–2020, respectively. Additionally, the Ontario study area forests displayed evidence of moisture stress, particularly in the summer months (June–August) from 1971–2010. The results showed forest cover will likely experience future losses and large fluctuations, putting the area's ecosystem at risk. Therefore, this method of monitoring forest degradation by integrating remote sensing tools could assist and improve climate change risk management, in particular by providing data on climate impacts on forest/vegetation.

Keywords: Forest disturbance and degradation, forest degradation indicators, Ontario forests, drought, LandTrendr.

4.1 Introduction

There is a deficiency of long-term, reliable, measured, and well-distributed meteorological variables in the study area of Ontario, Canada. These variables include, but are not limited to, the following: precipitation (P), ground runoff, evapotranspiration (ETo), soil moisture, including snow-water equivalent, and temperature (T). Many studies would benefit from having well-distributed (both in terms of time and space) and reliable estimates of these meteorological variables, which are crucial in many fields, including land use/land cover (LULC) planning and climate change studies. Unfortunately, in Ontario, Canada, this type of data is so incomplete (both temporally and spatially) that it has limited usefulness in ecosystem studies. This is a result of two factors: 1) the lack of resources to measure these variables over a large area and 2) the difficulty in measuring these variables reliably and accurately. Additionally, despite the relatively high density of meteorological stations in certain areas of the country, other parts (especially remote areas) have very few measurement stations (e.g., northern Ontario).

In fact, several studies have identified the importance of monitoring LC changes and vegetation dynamics for modelling the impacts of both climate and hydrological scenarios (Montandon and Small, 2008; Olmos-Trujillo et al., 2020). Widely used multi-spectral and multi-temporal images, which are sensed from Earth observation satellite systems, such as Landsat 8, can be employed in the study of LCLU changes over space and time, including mapping of vegetation and forest resources (Cui et al., 2013; Olmos-Trujillo et al., 2020). As vegetation responds to environmental characteristics, vegetation indices (VIs) have been found to be correlated with climatic (e.g., T and P) and other significant environmental variables (e.g., geomorphology and altitude) (De Keersmaecker et al., 2015; Olmos-Trujillo et al., 2020).

In order to monitor forest cover changes over time, forest degradation indicators (FDIs) computed from remote sensing imagery have proved particularly useful. Recent studies have shown the high efficiency of some indices (e.g., NDVI, NPP, LAI, and EVI) for the evaluation of spatial and temporal changes that

occur over forested areas across various scales (Prince et al., 2009; Fensholt et al., 2012; Eskandari Dameneh et al., 2021). Furthermore, FDIs are useful in the evaluation of ecosystem conditions (Zheng et al., 2018). The most common computed VIs (e.g., NDVI, NBR, EVI, TCT, and NPP) derived from multispectral imagery are also taken as indicators of vegetation and ecosystems conditions. These use the reflectance from vegetation and function under the same constraints as the source Landsat surface reflectance and MODIS products. They have their basis in the measurement of the radiation intensity of certain bands of the electromagnetic spectrum emitted, reflected, or transmitted by vegetation, as well as the estimated quality, development, and quantity of the vegetation (Vermote et al., 2016; Krakauer et al., 2017). Correlations between the NDVI and climate variables (e.g., T and P) are widely employed in many studies. However, other indicators (e.g., NPP, NBR, EVI, TCT) are also used as they are effective indicators of the response of different aspects of forest vegetation analysis in semi-arid/humid regions, such as the study area in central and northeastern Ontario, Canada.

While it is true that NDVI is one of the most commonly employed VIs, other indices have improved spectral surface response and attenuated other potentially confounding factors (e.g., soil, atmosphere, and light conditions) for exploration. In other studies, VIs have been compared based on a variety of different criteria (Fatiha et al., 2013; Olmos-Trujillo et al., 2020). For instance, Piedallullu et al. (2019) examined how climate change impacts vegetation cover; and Birtwistle et al. (2016) employed an alternate method for mapping the spatial and temporal variation of precipitation input along transient stream channels through NDVI taken from Landsat Thematic Mapper imagery.

In this study, the relationships between climate variables (T, P, ETo, and MAs) and FDI variables (NDVI, NPP, NBR, EVI, TCB, TCG and TCW) were identified and analyzed. In order to conduct this evaluation, remote sensing, GIS, and geostatistical tools were employed. The NDVI, NBR, EVI, TCG, TCB, and TCW were computed from multispectral Landsat imagery from 1972–2020, and NPP was obtained from

the MODIS platform for the series 2001–2020. A multiple linear regression model was developed, based on the most correlated FDIs and climate variables. The highest correlated variables were used to develop a multiple regression model for future prediction. In this study, all these indicators were combined for the creation of a novel, innovative, and more informative composite forest degradation indicator (CFDI) and the development of predictive models of future degradation (2050).

4.2 Background

According to the WMO, a climate normal is an arithmetic mean for a fixed 30-yr period that ends in a "10s" year (WMO, 1989). It is meant, in particular, to describe the mean seasonal cycle in terms of standard meteorological variables (e.g., P and T). Climate normals are commonly employed as baselines when characterizing and describing a particular meteorological dominant condition in an area. They are problematic, however, since, according to Arguez and Vose (2011), "climate normals are calculated retrospectively, but are often utilized prospectively." So, when climate normals are compared with weather forecasts, there is an assumption that they provide an estimate of the expected weather up to a certain date. However, this does not account for ongoing warming in temperatures. Additionally, when monthly datum is included in a temperature- and precipitation-normal calculation, no more than three consecutive days and no more than a total of five days can be missing for that particular month, which is known as the WMO "3 and 5 rule" (Braun et al., 2015).

As part of the process of creating spatial models, it is necessary to decide between more complete records and greater station density. So, in Canada, as in other parts of the world, one must make choices related to station data to be used for such models. Prior studies have resulted in the creation of a wide range of spatial models that cover various temporal and spatial extents (Price et al., 2000; Hutchinson et al., 2009; McKenney et al., 2011). While extensive testing has been part of these efforts, no study has revealed the relationship between monthly normals record completeness and model quality (Environment Canada,

1986a, b; Natural Resources Canada, 2011a, b). Instead, research in Canada has generally held that more station data are preferable, even when they are incomplete for the period, although many regions (such as central and northeastern Ontario) suffer from low station density.

As the Canadian climate network has temporal and spatial limitations, in this study the use of short- and long-term climate normals was optimized for the generation of spatial grids for monthly T_{max} and T_{min}, as well as precipitation normals. Daily minimum and maximum temperature and precipitation are variables employed in a wide variety of applications related to moisture balances. As a result, many experts have attempted to summarize them in various ways. For instance, daily minimum and maximum surface temperature can be accessed in the form of monthly averages for many 30-yr "normal" periods (e.g., 1971–2000, 1981–2010, 1991–2020).

Evapotranspiration (ETo) involves the transfer of water into the atmosphere due to evaporation from the surface of soil or water and from transpiration from plants. ETo returns to the atmosphere the moisture brought down by precipitation. It also has a close relationship with energy as measured indirectly by temperature (T_{max}, T_{min}, T_{mean}) and other climate parameters (e.g., wind velocity). The availability of moisture (MA) for plant growth, strictly speaking, should consider factors such as soil water holding capacity, infiltration rates, and ground water retention. However, for the purpose of a simple atmospheric balance, without incorporating the complexities of the soil and sub-soil media, it is calculated using a simple metric of cumulative P–ETo. The sum total of evaporation and transpiration processes, ETo depends on many conditions (e.g., radiation, energy, and wind) and vegetation factors (FAO-56; Allen et al., 1998). ETo links the energy and water cycles, which makes it an important contributor of water vapor to the demands of the atmosphere (Salam and Islam, 2020; Jerin et al., 2021). It is also a controlling factor in the water requirements for the growth of vegetation, amount of runoff from the soil, water loss from bodies of water and reservoirs, and water requirements for ecological sustainability. As a crucial player

in the water cycle, the estimation of ETo is fundamental to estimate a water balance, which in turn impacts every organism in ecosystems and, therefore, a wide range of sectors, including forestry (Ajjur and Al-Ghamdi, 2021).

WMO prescriptions are employed to obtain significant analyses for a full 30 years of observations (WMO, 2008); however, it must be considered that this would result in ETo underestimation stemming from ongoing climate change (Arguez and Vose, 2011). So, to ensure reliable and uniform ETo calculations, the FAO (FAO-56; Allen et al., 1998) has simplified the process through a parametrization of the Penman-Monteith equation (Trajkovic and Kolakovic, 2010; Valiantzas, 2013) resulting in the most reliable method for estimation of ETo. The calculation of potential ETo is done in standard conditions, that is to say, in an area that is completely vegetated, is uniform, and has a permanent water supply with shading. The estimation of ETo is influenced by several factors, such as T and P (Tabari and Talaee, 2014), vapor pressure (Irmak et al., 2006), relative humidity (Gong et al., 2006), and LU (Eichelmann et al., 2018). ETo is also reliant on a number of climatic parameters (e.g., solar radiation (SR), sunshine duration (SD), air temperature (T), relative humidity (RH), and wind speed (WS)), as well as their complex interactions, for which the ETo Calculator tool (FAO-56; Allen et al., 1998) offers default values for standard conditions that are identified and entered by the user of the tool.

Fully understanding and predicting how forests respond to disturbances is crucial for the management of forests and source-water areas, such as the study area, with a relatively high ETo regime. There is a particular and pressing concern in areas where a combination of factors (e.g., global warming and past management) has had negative impacts on the environment, including the following: (i) intense and wide-ranging wildfires (McKenzie et al., 2004; Westerling, 2006; North et al., 2015a), (ii) high incidence of forest mortality related to drought (Allen et al., 2010; Anderegg et al., 2015), and (iii) a reduction in runoff (Barnett et al., 2005; Goulden and Bales, 2014). The response of runoff to ETo could be greater than its

response to disturbances and interannual changes in temperature and precipitation, depending on many other contextual circumstances. Therefore, evapotranspiration is an important metric for determining changes in the moisture balance which will determine moisture availability for forested areas. Many studies in different ecosystems (Jiapaer et al., 2015; Ren et al., 2020) over recent decades have shown how vegetation dynamic changes are closely related to climatic variation. So, the analysis and prediction of vegetation dynamic changes, especially with the increasing influence of climate change, are key to understanding its significance in determining carbon sequestration in vegetation biomass, litter, and soil organic matter (Fu et al., 2010; Zhou et al., 2020).

Since forests are dynamic systems, they can react to biomass reductions in a variety of ways in relation to their water usage (Tague et al., 2018). With the emergence of high-confidence spatial-ETo estimates that are driven by a strong relationship between vegetation greenness estimates (e.g., NDVI, and others, derived from multispectral satellite images, and ground-point measurements of ETo in various ecosystems; Goulden et al., 2012; Goulden and Bales, 2019), moisture balance can be estimated with high spatial resolution across forested landscapes. For forest management, this ETo mapping tool allows the user to estimate the ETo changes resulting from previous wildfires and fuel treatments (Roche et al., 2018), as well as predict future changes from treatments, disturbances, and degradation (Roche et al., 2020). Additionally, ETo, which functions as another basic part of the hydrological cycle and is related to water and energy exchange between the Earth's surface and its atmosphere, has the capacity to determine the energy and water transport in the crucial soil-vegetation-atmosphere system (Shi et al., 2017; Zhou et al., 2020). Many studies have also been conducted on how ETo impacts vegetation dynamic changes (Sun et al., 2012; Zhang et al., 2019; Zhou et al., 2020). As a thorough comprehension of ETo has the potential to assist in ecosystem services and water resource management, it is worthwhile to

explore the influence of ETo on NDVI, including other VIs, to ensure vegetation changes are monitored in a dynamic way (Zhou et al., 2020).

In every region, climate is distinctly characterized by precipitation and temperature. In regions where the ETo exceeds P, there is a moisture deficit, which falls into a variety of classes. For instance, the South Saskatchewan River Basin (SSRB) is classified as semi-arid, dry-subhumid, sub-humid, or humid according to the definitions below (Marchildon, 2009):

Semi-arid: P is less than 1/2 ETo (P<0.5 ETo)

Dry-subhumid: P ranges from 0.5 to 0.6 ETo

Subhumid: P ranges from 0.6 to 1.0 ETo

Humid: P is greater than ETo (P>ETo) (MA_1= P-ETo is positive)

Drought, which can be defined as a deficiency of P relative to ETo over an extended period (a season or more) leading to a water shortage with adverse impacts on living organisms, is a crucial threat to the future functioning of the boreal forest ecosystems. Evaporative demand, which can be represented by the equation to estimate ETo, is a result of the combined effects of radiative heating and atmospheric humidity deficit (e.g., Monteith and Unsworth, 2008). The estimation of ETo, which can be defined as the estimated water vapor loss from vegetation that is healthy and productive and growing in soil with adequate moisture, is obtained with simple equations that employ either a combination of long-term monthly temperature and solar radiation-related data (Jensen-Haise method; see Hogg, 1994) or only temperature data (simplified Penman-Monteith method, Hogg, 1997).

Hogg (1997) detailed the development of a regional climate moisture index (CMI), which is equivalent to moisture availability (MA). The MA is calculated as P minus ETo, wherein P is the (daily or monthly) average precipitation (including rain and snow melt) and ETo is the daily or monthly ETo, both of which are expressed in cm or mm of water per year. If the MA values are positive, this means conditions are

moist or wet and that there is enough precipitation to sustain a forest with a closed canopy. On the other hand, negative MA values are an indication of dry conditions that, in the best-case scenario, have the ability to sustain discontinuous forests. The MA values are positive (P>ETo) in the boreal forest, whose landscape is primarily composed of conifers and peatlands, and water runoff on an annual basis is sufficient to ensure and maintain ongoing stream flow and lake levels that are stable (Campbell et al., 1994). In contrast, where MA values are negative (p<ETo), it is rare to find conifers or peatlands.

Based on a moisture balance model, the length of growing period (LGP) quantifies the length of time during which sufficient moisture is available to sustain the growth of vegetation. It can be defined as the number of days when the temperature allows for plant growth and soil moisture supply is greater than half ETo (P>0.5ETo). When considered in combination with common precipitation distribution patterns, such a definition accounts for moisture requirements of vegetation in the early growth stages and at maturation/ripening that are lower than full ETo, and that in between the requirements are close to full ETo. Growth periods during which there is a sub-period when precipitation exceeds full potential ETo are labelled normal LGPs, as opposed to intermediate LGPs, where precipitation is greater than 0.5 ETo but does not attain full ETo. These different LGP types are considered in the assessment of vegetation yield potentials (FAO, 1978–81). Basic climate data are employed in the calculation of derived climate variables, daily moisture balance, ETo, T zones, as well as LGP pattern and length. The LGP analysis for potential circumstances that might imply increased CO_2 concentrations takes into consideration the slowing down of transpiration that results from the partial closure of leaves' stomata, by means of the modification of the canopy resistance term for the ETo formula. Determining the potential changes in LGP of vegetation is one means of assessing the impact of climate change as studies have shown how plants' LGP can be significantly increased or decreased as a result of climate change. In order to develop a detailed picture of the effects of climate shifts on the growth and development of plants (which can be

positive or negative), LGP can either be combined with or compared to other elements of projected climate change (e.g., changes in T).

In terms of interpolation techniques for climate data, kriging is advantageous due to its greater accuracy since the weight selections are not arbitrary; rather, these are chosen depending on the behavior of the semi-variogram function over the space (Kitanidis, 1997). Kriging also furnishes the means to make an evaluation of the magnitude of estimation error (Kitanidis, 1997). This technique fits the theoretical variogram to the experimental variogram by automatically adjusting the model parameters obtained from simulation-based optimization (Krivoruchko, 2012a). Kriging with an external drift appears to have greater sensitivity with respect to the quality of radar data (Berndt et al., 2014). EBK has the following positive characteristics: the specification of the prior distributions for the model parameters is not required; it permits the non-stationarity of moderate local and large global data; and locally, it transforms data to the Gaussian distribution, if necessary. EBK allows for varying measurements of error; employs covariates in order to enable co-regionalization estimation, if needed. Moreover, EBK works faster and produces outputs that are reliable resting on default parameters. The use of EBK in analysis and spatial prediction (Esri, 2019a, c; Krivoruchko, 2012; Krivoruchko and Frączek, 2015) has been conducted through software applications. Given its features, this study used EBK for the interpolation of the climate variables within multi-climate datasets of the study area by first selecting the best-fitting EBK semi-variogram model to understand the spatial and temporal variability of the research zone daily multi-climate variables.

The overall objective of this research was to ascertain how changing environmental conditions impact forest degradation in the study area, by conducting an analysis of the vegetation dynamic changes during the period from 1972–2020, and projecting changes in 2050. Accounting for the effects of T, P, and ETo on vegetation from the perspective of water deficit, multiple linear regression analysis was employed as

the basis for developing predictive models. With these predictive models and the meteorological data from general circulation models (GCMs), specifically from CMIP5 experiments, the future dynamic changes of forest degradation in the study area were projected, assuming the relationships found with historic data will remain unmuted and can be projected into the future. In particular, this work aimed (i) to use the established methodological framework to assess the trends in forest cover change in the region during the longest period possible (e.g., 1972–2020), and the relationship between the trends of forest cover change and their underlying causative factors, particularly their relationship to climate change, and (ii) to find out the dependence of disturbance or degradation on or at least its degree of correlation with climate variables.

4.3 Materials and methods

Daily minimum and maximum temperatures (T_{min}, T_{max}), and precipitation (P) are data frequently used for climate assessments. Historical climate (daily and monthly) data for the period from 1972–2020 were obtained from the official climate data sources of Environment and Natural Resources Canada. In addition, historical Canadian climate normals data were acquired from Environment and Climate Change Canada. Two 30-year climate normals datasets (1971–2000 and 1981–2010) were collected from 51 weather stations inside and surrounding the study area, including a few stations in the Province of Quebec, in order to acquire accurate data within the rectangular perimeter of the research zone. These 51 stations (Figure 4.3) meet the WMO "3 and 5 rule" described above (Braun et al., 2015). As a large portion of the study area is uninhabited, it has very few weather stations and sparse observed climate data. Thus, only precipitation and temperature data are available from observed meteorological datasets over the whole of the study area.

For this study, it was determined that at least 20 years of data must be available for a station to be included in the analysis. This goal was to include the largest number of meteorological stations possible for maximum spatial coverage and to capture the pattern of temporal variability. Stations were distributed

uniformly across the study area with a slightly lower station density in the north (Figure 4.3). To mitigate the boundary edge effect during the spatial interpolation of each variable, some station data were collected from the neighboring province of Quebec. Station location and elevation data (which ranged between 69m and 490m) were also recorded. The period of 1971–2000 was chosen since data from the maximum number of meteorological stations was available in Ontario, even if the records are incomplete for the period being studied (Environment Canada, 1986a, b; Natural Resources Canada, 2011a, b), particularly in a region where station density is relatively low.

Seven different VIs derived in this study as FDIs, such as NDVI, NPP, NBR, EVI, TCG, TCB, and TCW, were calculated using ENVI^{TM} 5.3 software. These were employed to analyze the relationship between FDI variables and climate variables (T, P, ETo, MA_1, and MA_2) from 1971–2020. The selected indicators were calculated from the Landsat multispectral imagery, including moderate resolution imaging spectroradiometer (MODIS) images. Since Landsat images have a spatial resolution of 30m, the MODIS products (MODIS NPP, etc.) were resampled to 30m, to match the spatial resolution of the Landsat imagery. Pearson correlation coefficients (r) between FDIs and climate variables were investigated. The correlation matrix is shown in Table 4.5. All FDIs were combined into a CFDI variable to obtain the most informative FDI. This was achieved using an image fusion technique based on PCA within the ENVI computing environment. In order to ensure consistency with the indicators' data run length (i.e., from 1972–2020), meteorological data from 1971–2020 were also selected to match the dates. Unfortunately, 1991–2020 climate normals data have not yet been released, so a manual calculation was used on the 2011–2020 daily historical data.

ETo refers to the atmospheric demand for evaporation and transpiration from a reference static vegetation surface when there are no water limitations and is calculated with the Penman-Monteith equation furnished by the Food and Agriculture Organization of the United Nations (FAO) (Yao et al., 2018; Allen

et al., 1998). This method requires the following data: T_{max} and T_{min}, daily average dew-point T (or the equivalent, which is vapour pressure or vapour pressure deficit), downward shortwave radiation, and wind speed. The software tool (ETo Calculator) in which the ETo method was implemented is capable of providing adequate default values for difficult-to-obtain variables (e.g., vapour pressure, radiation, etc.).

Climate normals data for the two historical periods (1971–2000 and 1981–2010) were compiled for ETo calculation, according to the guidelines from The ETo Calculator Reference Manual Version 3.2 (FAO, 2012). Some essential intermediate formulas and procedures were necessary to solve the equation selected to calculate reference potential ETo based on the Penman-Monteith method (Hargreaves and Samani, 1982; Allen et al., 1998). The observed ETo was unavailable in the study area. Therefore, the Penman-Monteith (FAO-56) method was used to calculate the ETo, which is recommended as the ideal technique by the FAO when observed ETo is not readily available (Allen et al., 1998; Feng et al., 2017), something that has proven true for various climate settings and time-step calculations lacking local calibration (Valiantzas, 2013; Feng et al., 2017; Islam et al., 2019). The estimation of ETo from the Penman-Monteith simplified method is obtained from equation 4.1 below:

$$\mathrm{ETo} = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{(T+273)} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (4.1)$$

Where ETo indicates the reference potential evapotranspiration ($mmday^{-1}$); R_n represents net radiation ($MJm^{-2}day^{-1}$); G represents soil heat flux density ($MJm^{-2}day^{-1}$); T indicates T_{mean} (^0C); e_s stands for mean saturation vapor pressure at air T (KP_a); e_a represents actual vapor pressure derived from RH mean (KP_a); Δ represents the slope of saturation vapor pressure curve at T_{mean} (kPa^0C^{-1}); Υ indicates theoretical psychrometric constant ($KP_a{}^0C^{-1}$); and U_2 represents wind speed at 2m above the ground (ms^{-1}).

The FAO-56 Penman-Monteith equation was also employed to calculate daily soil moisture levels. To determine ETo rates on a surface, this equation makes use of the following four climatic parameters averaged over each day: wind speed (WS) (ms^{-1}), solar radiation (SR) ($MJm^{-2}day^{-1}$), air T (^0C), and dew-

point T (^{0}C) (Zontarelli et al., 2010). On days with copious precipitation, the moisture lost through ETo is replenished, which will raise the moisture content of the soil. Reference ETo rates (mm) that have been estimated by the Penman-Monteith method rely on the assumption that the available water is unlimited, which negates the moisture-retention capacity of different soils. Alternately, physical-based approaches to ETo scaling necessitate a number of attributes (spatially resolved radiation, humidity, temperature, wind speed, etc.) that display substantial variation over small distances, resulting in highly uncertain extrapolations of mountainous meteorological conditions on a fine scale (Goulden et al., 2012).

The climate moisture index (CMI) shows a good correlation with tree growth in the ecosystems of temperate and boreal forests (Berner et al., 2017; Hogg et al., 2013). CMI (MA) (cm/month^{-1}) (mmday^{-1}) values were calculated as precipitation (P) (mm/day^{-1}) minus ETo (mm/day^{-1}) (P–ETo) (based on Hogg, 1997). ETo, which was estimated in mm/day^{-1} and included the expected rate of water vapor loss to the atmosphere, was calculated from average daily T_{min} and T_{max} (^{0}C) for each month (monthly average of 40 years from 1971–2010) of climate normals data using the FAO-56 (Allen et al., 1998) and "ETo Calculator" version 3.2. In this study, a 12-month window (January to December) was employed for MA (P–ETo) calculation. Then, the MA was calculated using the daily average of 12 months for one station. Positive MA values indicate wet or moist conditions, whereas negative ones denote dry conditions. The ETo, MA_1, and MA_2 were calculated in the same way for all 51 stations, where MA_1=P-ETo and MA_2=P-0.5ETo. In this study, daily results for ETo were used only after comparing the performance of the calculated daily and monthly calculations. The P–ETo difference was also employed to calculate the soil-moisture deficit. Statistical modelling (as well as interpolation methods) were conducted using spreadsheet in ArcGIS^{TM} and ArcGIS Pro^{TM} software for the analysis.

Derived 40-year average daily raster/gridded maps of MA_1 and MA_2 involve an estimate of the moisture amount available for runoff or storage recharge (i.e., excess) or the amount of moisture required from

storage (i.e., deficit) to maintain estimated annual/daily ETo. The reason for computing two values of MA was to attempt to understand the relationship between FDIs and two levels of moisture deficit or availability, whereby MA_1 is more restrictive than MA_2, which is more permissive, given that only half of the ETo is subtracted. The latter MA is employed in global assessments (FAO, 1978–81) to calculate the LGP for vegetation.

The linear trend for a time series analysis is computed with the least square method, which is represented by linear regression equation (LRE) 4.2:

$$Y_n(t) = a_0 + a_1 t \quad (4.2)$$

Where the slope a1 is the estimated trend, and a1X10 is called the climate tendency rate, which presents its change rate per decade. A negative value of a1 signifies a negative trend, and a positive value is indicative of a positive trend.

Multiple linear regression (MLR)

MLR, which employs least squares, involves minimizing the sum of the squares of the deviations of the observed response from the fitted response (Naoum and Tsanis, 2003; Zhang et al., 2018). With CFDI being the dependent variable, a model was constructed involving the environmental variables and their corresponding parameters. The model's general form is presented below:

$$CFDI = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_5 X_5 \quad (4.3)$$

Where CFDI is the degradation indicator; X1–X5 are T, P, ETo, MA_1, and MA_2, respectively. Furthermore, β_0—β_5 are the corresponding parameters of the regression function. Stepwise regression was employed to avoid multicollinearity and LRE overfitting. For the whole time period, the significant correlation between VIs/FDIs and climate variables was found to be at the 0.05 significance level, while the other four factors had no significance at the 0.1 level. Thus, the regression relationship was founded based on CFDI, P, MA_1, and MA_2.

The equations to predict the CFDI values from climate variables were constructed on the basis of the MLR model. For each grid, the CFDI values can be expressed as a function of T, P, ETo, MA_1, and MA_2. Therefore, using the values of the indicators from 1972–2020 as the dependent variables and the climate variables during the period from 1971–2010 as the independent variables, the regression equation could be expressed in the following manner:

$$CFDI = \beta_0 + \beta_1 T + \beta_2 P + \beta_3 ETo + \beta_4 MA_1 + \beta_5 MA_2 + \varepsilon \quad (4.4)$$

Where β_0=intercept, β_1–β_5 are the coefficients (slopes) of the multiple regression equation (independent variables), and ε is a constant/residual.

In order to provide spatially explicit models of variability, the regression equations were built in the raster calculator of GIS software in order to generate raster maps of elevation, T, P, ETo, MA_1, and MA_2 meteorological variables over the entire study area at 30m spatial resolution. Additionally, given the fact that the data from the weather stations and the Landsat products came from different sources with different reference projections for the environment under study, ArcGIS™ was employed to adapt them to a consistent geographic coordinate system. During this process, the weather station locations were converted to the point geometry feature class with the ArcGIS tool. Moreover, mean meteorological attributes (T, P, ETo, MA_1, MA_2) were spatially interpolated to create several weather layers. The method chosen for spatial interpolation and spatial resolution of the meteorological stations is significant because it could have an effect on the ETo spatial pattern (Tang et al., 2013; Feng et al., 2017). For this study, the EBK method was shown to have the lowest mean error when compared to other interpolation methods; therefore, EBK was used in the creation of T, P, ETo, MA_1, and MA_2 raster/gridded maps in the ArcGIS™ environment. An output cell size was chosen to match the spatial resolution (30mX30m) of FDIs. The final output consists of daily average maps for the study period from 1971–2010 using the daily average climate normal data from that period. Additionally, mean annual NDVI, NPP, and CFDI products series

and mean annual P, MA_1, and MA_2 gridded maps were overlapped for statistical analysis of the forest degradation trend, resulting from analysis with the LandTrendr algorithm (details in Chapter 3). Finally, least squares regression model fit was employed to examine the relationships between the climate variables and FDIs.

The goal of this study was to produce the most accurate gridded dataset of T, P, ETo, MA_1, and MA_2 for the study area based on measured climate normals data from 1971–2010 so that relationships between climate and forest degradation indicators could be explored. In order to create useful gridded meteorological datasets from the station data, many different methods exist to interpolate such data. However, P and T are the only meteorological variables measured throughout the study area with enough spatial and temporal density to guarantee meaningful results from spatial interpolation. The choice of an accurate method of optimizing the interpolated values could be subjective. However, this is a crucial step in producing accurate distribution maps of phenomena that are in and of themselves continuous, such as T and P (Addis et al., 2016; Wright, 2020).

Various semi-variogram models were considered to model the spatial variability of meteorological variables (i.e., Spherical, Circular, Exponential, Gaussian, K-Bessel, J-Bessel, and Hole Effect). They were scrutinized for their significant effect on predicting unknown climate variable values. Cross-validation focusing range parameter estimation was used in this study (Ali et al., 2021) to optimize the semi-variogram models and associated parameters, including nugget, range, and sill (Hunter et al., 2020). Li and Heap (2011) reviewed 53 studies and 72 interpolation methods/sub-methods frequently used in environmental sciences. Their results revealed that overall, kriging methods display better performance than non-geostatistical methods at annual or monthly time scales (Li and Heap, 2008; Li et al., 2011; Bolivar, 2020). In order to ascertain the best-performing method, the choice of parameters/models (e.g., semi-variogram models) among other criteria is significant in the accuracy of results. Unfortunately,

finding the most robust method for a certain study is generally a long process and affected by decisions on the acceptable prediction error for a particular location and variable of interest (e.g., soil properties).

To determine the optimal methods for spatial prediction, standardized indicators were considered: as MSE represents the standardized error average, it should come close to zero, in addition to RMSE, which denotes the distance the expected values deviate from the real value, with one representing the ideal match between predicted and measured. For the interpolation process, the value with the lowest residual was deemed to have the best reliability.

The interpolation of daily climate/multi-climate data, which was done from the irregularly spaced station locations to the nodes present on a regularly spaced 30m grid, was completed with the use of the EBK spatial interpolation algorithm. Despite the exploration of other methods (e.g., kriging, thin plate splines), the low topographic complexity of and high station density in the study area showed relatively high-quality results with the EBK interpolator. The present study employed EBK in an ArcGIS Pro environment to confirm its robustness and also visualize the spatial variability of the meteorological variables (which can also create a relationship with FDIs, including CFDI). As previously stated, EBK automates the most complicated feature, which is the semi-variogram modelling, in the construction of a valid kriging model, which is also the aspect of spatial interpolation that is the most challenging within the geostatistical methods (Gupta et al., 2017). EBK was also selected since it furnishes more accurate spatial estimation of the meteorological variables (T, P, ETo, MA_1, and MA_2). Additionally, its performance is extremely high in goodness-of-fit tests of r (correlation coefficient) and RMSE compared to other kriging methods (Gupta et al., 2017).

In this study, the spatial interpolation of the datasets was performed for the average of 40 years (1971–2010) using the ordinary kriging (OK) and EBK geostatistical modelling techniques. Prior to their

application, the collected datasets were assessed for their regularity and correctness. Subsequently, attributes that were relevant to the data locations were assigned in the GIS platform.

Cross-validation assesses the results of a statistical analysis concerning an independent dataset (Chen and Liu, 2012). It helps in the selection of the optimal geostatistical model that furnishes the best predictions over the space (Tveito et al., 2008). It is primarily employed to estimate a predictive model's accuracy and to establish the objective to gain a high level of prediction accuracy. Additionally, it has been widely used to evaluate the accuracy of geostatistical modelling techniques. There are two goodness-of-fit criteria (i.e., Pearson correlation coefficient (r) and coefficient of determination (R^2)) that are generally employed with the cross-validation technique for determining the best-fit model in hydrological studies.

For spatial data analysis, large datasets can generally be divided with the result that the selected EBK model can be successfully fitted to the data in each subset. The EBK's default model is intrinsic random function kriging (IRFK) with power semi-variogram. This effectively predicts locally varying data and generally functions well with dependent stationary data. In order to decrease computation time, the default model employs only the 10 nearest observations. Owing to their smaller number of model parameters, the IFRK models work much faster than other options. Therefore, as one might expect, the longest running time is seen in the model with the largest number of parameters, the K-Bessel with linear trend models, and data transformation (which was used for this study as the most flexible and accurate model).

For those using Geostatistical Analyst, there will be a difference in the semi-variogram modelling in the Geostatistical Wizard. As shown in Fig. 4.1, the semi-variogram model appears to have a zero "nugget effect" parameter in cases where individual measurement errors are furnished as EBK-model input, while the data variation at zero distance between the data locations is sustainable in instances where measurement errors are unavailable for geostatistical modelling. It should be noted that the nugget tab is missing in the former case.

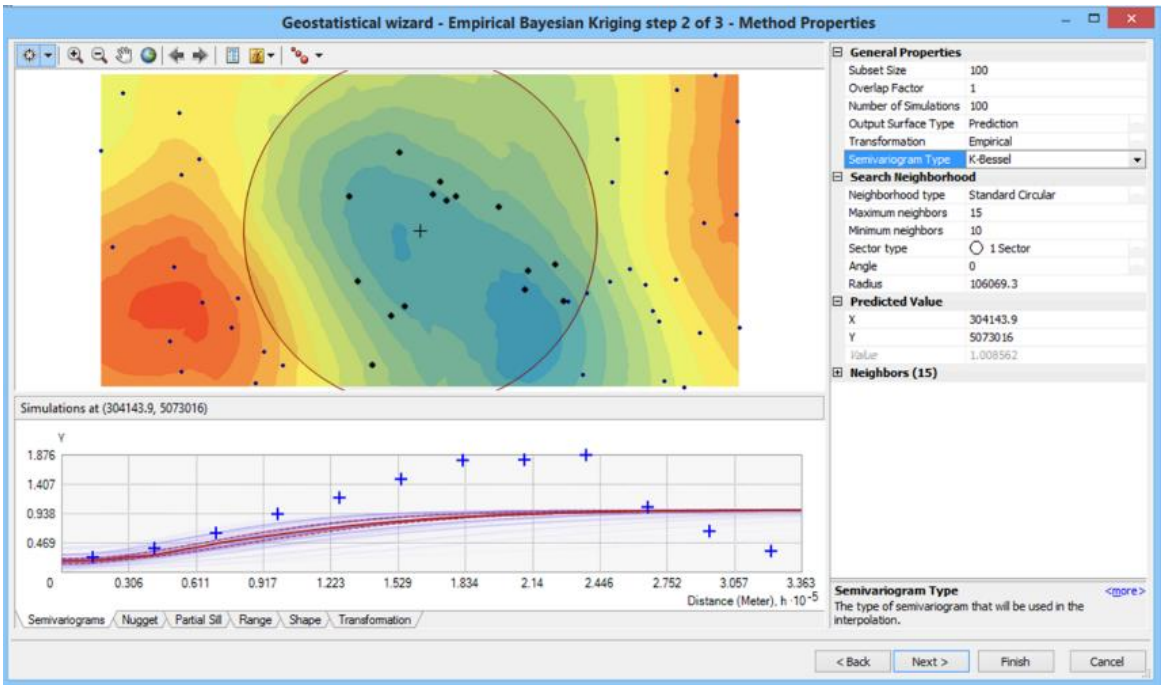


Figure 4. 1: Screenshot of the semivariogram model fitted to data, including other tabs

User-defined subsets are chosen or the parameters for the creation of new subsets are accepted or defined. Next, the default EBK model can be run by the user. Alternately, the generalized covariance model, the neighborhood search parameter, simulation number, and the model's output before running the model can be altered. There is a geostatistical layer created by the software, which involves a data structure that stores both the source-data path and the parameters applied for interpolating values. The geostatistical layer is employed for map creation (quantile, prediction and prediction standard error, and probability of

exceeding a specified threshold value), and to evaluate/update models already created in the Geostatistical Wizard. It is also possible to produce model diagnostics for point locations and separate data-subset summaries. However, rather than straight lines, the data transformation to Gaussian distribution and the chordal distance metrics are employed automatically on occasions where the linear mixed model (LMM) is used without covariates and the data coordinates are latitude and longitude, respectively.

Various cross-validation parameters are employed to assess the performance of the different interpolation methods. This is a common statistical approach used to examine the exactness of interpolated data. Error determination is conducted through the estimation of the difference between each observed (measured) value and its predicted counterpart. These error tests should function as reasonable evaluation estimators that can be used to compare different interpolation models, thereby acting as the foundation for the method selection process. Generally, studies employ only one or two parameters to choose the most appropriate interpolation method. RMSE and R^2 are the most commonly used for this process (Guenni et al., 1998; Lai et al., 2019). In applied geostatistics literature, it is generally recommended to employ a neighborhood with 8–16 observations (Chilès and Delfiner, 1999). The EBK data subsets are limited to 1000 observations, all of which can be used for predictions. For unsampled locations in EBK, however, the maximum number of allowable observations for prediction is 64.

The mean value of climate variables for 1971–2000 and 1981–2010 (average of two climate normal periods) were employed as baseline conditions in the current study. Mean climate variables for observation point were interpolated monthly in GIS (interpolated by EBK) to create a 30 X 30m grid (for a full account of the methodological workflow for this chapter, see Appendix 4).

4.4 Results and discussion

Two sets of 30-year climate normals data (1971–2000 and 1981–2010) were collected from 51 weather stations inside and surrounding the study area, including a few stations from the neighboring province of

Quebec, in order to acquire accurate data within the rectangular perimeter of the research zone. The data were compiled and prepared to calculate the ETo using the ETo Calculator tool (FAO-56), MA_1, and MA_2 for all 51 weather stations. Table 4.1 shows a fraction of the calculated summary with one set of 30-year average climate normals data for the Combermere meteorological station.

Table 4. 1: Summary of the 30-year climate normals (1981–2010) data for the Combermere station in Ontario (calculated as daily and monthly averages)

Months	Daily Average			Monthly Average		Daily Average		Monthly Average	
	T	P	ETo	P	ETo	MA_1	MA_2	MA_1	MA_2
January	-11.6	2.0	0.6	63.3	18.6	1.442	1.742	44.700	54.000
February	-9.4	1.8	0.9	49.5	25.2	0.868	1.318	24.300	36.900
March	-3.8	1.9	1.5	58.9	46.5	0.400	1.150	12.400	35.650
April	4.4	2.2	2.5	66.8	75.0	-0.273	0.977	-8.200	29.300
May	11.5	2.5	3.9	78.2	120.9	-1.377	0.573	-42.700	17.750
June	16.8	2.7	4.8	81.0	144.0	-2.100	0.300	-63.000	9.000
July	19.4	2.4	5.0	75.4	155.0	-2.568	-0.068	-79.600	-2.100
August	18.1	2.2	4.3	68.9	133.3	-2.077	0.073	-64.400	2.250
September	13.4	2.9	3.1	85.5	93.0	-0.250	1.300	-7.500	39.000
October	6.7	2.7	1.8	85.0	55.8	0.942	1.842	29.200	57.100
November	0.4	2.6	1.0	79.0	30.0	1.633	2.133	49.000	64.000
December	-6.8	2.1	0.7	65.5	21.7	1.413	1.763	43.800	54.650
Average	**4.9**	**2.3**	**2.5**	**71.4**	**76.6**	**-0.162**	**1.092**	**-5.167**	**33.125**
Total				**857**	**919**			**-62**	**397.5**

The figures for Combermere station indicated a moisture deficit from April to September in this very demanding area in terms of moisture availability, particularly during the summer months. In contrast, when precipitation is compared to 0.5ETo (MA_2; as in most moisture balances used for estimation of the LGP, e.g., FAO framework, 1978–81), only in the mid-summer month (e.g., July) there appears to be a moisture deficit in the study region. Yearly average results also revealed a moisture deficit (daily and monthly average deficits were 0.162mm and 5.167mm, respectively in the Combermere station area) (Table 4.1), even though for some stations, an increase in precipitation was also accompanied by a moisture deficit due to the rise in temperature resulting in higher evapotranspiration.

Table 4. 2: Calculated results for comparison of the daily and monthly ETo, MA_1, and MA_2 for 21 meteorological stations

				Daily Average					Monthly Average			
Stations	**Lat**	**Long**	**Elev**	**T**	**P**	**ETo**	**MA_1**	**MA_2**	**P**	**ETo**	**MA_1**	**MA_2**
Arnprior Grandon	45.42	-76.37	107	5.7	2.2	2.4	-0.221	0.991	67.1	74.1	-6.942	30.100
Beatrice	45.13	-79.4	297.2	4.8	3.3	2.6	0.671	1.975	99.8	79.6	20.183	59.996
Chalk River	46.05	-77.37	121.9	5.6	2.4	2.5	-0.112	1.121	71.6	75.3	-3.692	33.963
Chelsea QC	45.52	-75.78	112.5	5.9	2.7	2.4	0.283	1.483	81.7	73.3	8.425	45.071
Combermere	45.37	-77.62	287	4.9	2.3	2.5	-0.162	1.092	71.4	76.6	-5.167	33.125
Dorset Moe	45.22	-78.93	323.1	5.0	2.9	2.5	0.398	1.648	88.2	76.3	11.858	50.029
Drummond Centre	45.03	-76.25	145	6.1	2.4	2.7	-0.267	1.066	73.0	81.4	-8.408	32.308
Dwight	45.38	-78.9	404	4.3	3.2	2.4	0.880	2.063	98.8	72.3	26.550	62.675
Haliburton	45.03	-78.53	330	5.0	2.9	2.7	0.257	1.599	89.5	81.9	7.542	48.504
Huntsville	45.35	-79.17	320	5.5	2.8	2.4	0.415	1.623	86.2	73.8	12.383	49.275
Luskville QC	45.53	-76.05	68.9	5.6	2.6	2.7	-0.084	1.245	78.4	81.2	-2.792	37.796
Madawaska	45.5	-77.98	316.4	4.2	2.4	2.7	-0.358	0.997	71.6	82.7	-11.117	30.242
North Bay	46.36	-79.42	370.3	4.2	2.9	2.2	0.709	1.784	87.1	65.7	21.400	54.229
Ottawa CDA	45.38	-75.72	79.2	6.6	2.5	2.4	0.108	1.312	76.6	73.5	3.083	39.854
Ottawa-McDonald	45.32	-75.67	114	6.4	2.6	2.5	0.133	1.358	78.6	74.8	3.825	41.229
Powassan	46.12	-79.25	274.3	4.4	2.6	2.4	0.214	1.406	79.2	72.8	6.367	42.758
Rapid D Joachims QC	46.2	-77.7	137.2	5.2	2.4	2.6	-0.135	1.157	74.6	78.9	-4.317	35.125
Renfrew	45.48	-76.7	129.5	5.0	2.2	2.7	-0.489	0.844	66.3	81.5	-15.175	25.558
Shawville QC	45.62	-76.47	167.6	5.1	2.5	2.5	-0.061	1.197	74.8	76.9	-2.092	36.338
Sheenboro QC	45.97	-77.25	137.2	5.0	2.3	2.6	-0.309	1.012	71.1	80.7	-9.567	30.763
Wright QC	46.07	-76.05	141.7	4.7	2.6	2.5	0.047	1.309	78.3	77.1	1.208	39.767

Monthly average temperature data was applied as input, which is a prerequisite for the ETo Calculator tool. Canadian climate normals precipitation data were acquired as monthly totals. To determine the best performance (between monthly and daily average calculation for EBK application), the predicted scatter plots (best fit) were compared using monthly and daily average data. Daily average showed better performance (Figure 4.2a).

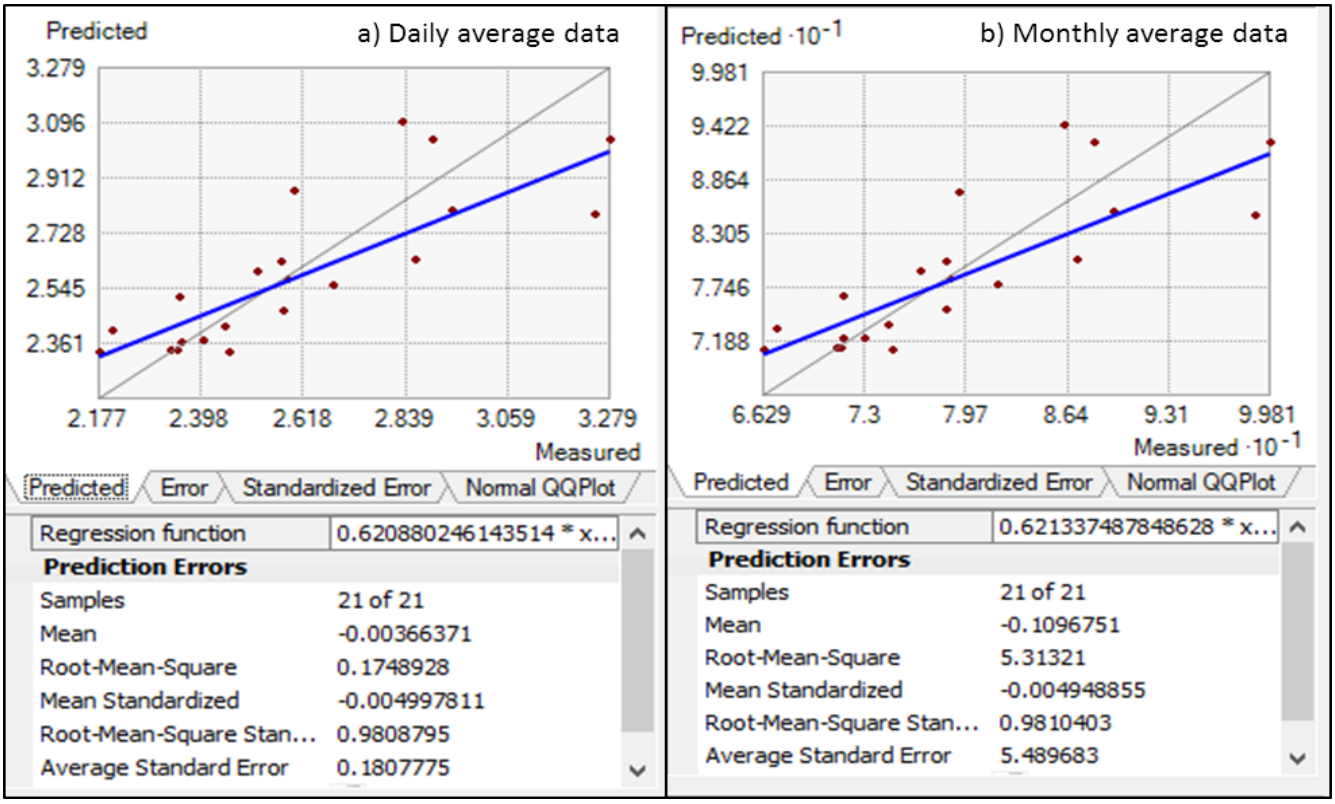


Figure 4. 2 (a-b): The predicted best fit model with prediction errors of (a) daily average precipitation data and (b) monthly average precipitation data (Note that daily geo-statistical layer maps are smoother than monthly)

4.4.1 Results of calculations for 51 stations

Daily averages of T, P, ETo, MA_1, and MA_2 were calculated for 51 meteorological stations for the climate normals of two 30-yr periods: 1971–2000 and 1981–2010 (Table 4.3). The results of the two sets of calculations were combined into one as a 40-yr average (1971–2010).

Table 4. 3: Daily averages of 40 years (1971–2000 and 1981–2010) for T, P, ETo, MA_1, and MA_2 were calculated for 51 climate stations (compiled/summary table)

Stations	**Lat**	**Long**	**Elev**	**T**	**P**	**ETo**	**MA_1**	**MA_2**
Angers QC	45.55	-75.55	91	4.9	2.7	2.5	0.203	1.488
Arnprior Grandon	45.42	-76.37	107	5.6	2.1	2.4	-0.223	0.986
Beatrice	45.13	-79.4	297.2	4.6	3.2	2.5	0.692	1.986
Chalk River	46.05	-77.37	121.9	5.4	2.3	2.4	-0.113	1.12
Chelsea QC	45.52	-75.78	112.5	5.7	2.6	2.3	0.35	1.521
Cheneville QC	45.9	-75.08	222.5	4.2	3	2.4	0.514	1.76
Combermere	45.37	-77.62	287	4.7	2.3	2.4	-0.133	1.114
Dorset Moe	45.22	-78.93	323.1	4.9	2.9	2.5	0.409	1.659
Dunchurch	45.62	-79.88	268.2	4.7	3	2.5	0.5	1.778
Dwight	45.38	-78.9	404	4.1	3.2	2.3	0.883	2.062
Haliburton	45.03	-78.53	330	4.9	2.8	2.5	0.292	1.572
Huntsville	45.35	-79.17	320	5.5	2.8	2.3	0.432	1.63
Kemptville	45	-75.63	99.4	6	2.6	2.4	0.14	1.382
Luskville QC	45.53	-76.05	68.9	5.5	2.5	2.6	-0.127	1.201
Madawaska	45.5	-77.98	316.4	4.1	2.3	2.7	-0.375	0.976
Minden	44.93	-78.73	274.3	5.4	2.8	2.7	0.142	1.508
Mont Laurier QC	46.57	-75.55	244	3.7	2.7	2.4	0.323	1.537
Muskoka	44.97	-79.3	281.9	5	3	2.5	0.511	1.763
Nomininguе QC	46.4	-75.08	274	3.4	2.9	2.3	0.569	1.761
North Bay	46.36	-79.42	370.3	4	2.8	2.1	0.662	1.735
Ottawa CDA	45.38	-75.72	79.2	6.4	2.5	2.3	0.125	1.317
Ottawa McDonald	45.32	-75.67	114	6.1	2.5	2.4	0.157	1.369
Powassan	46.12	-79.25	274.3	4.3	2.5	2.4	0.165	1.372
Rapid Des Joachims QC	46.2	-77.7	137.2	4.9	2.4	2.5	-0.103	1.181
Renfrew	45.48	-76.7	129.5	4.9	2.1	2.6	-0.454	0.872
Russell	45.26	-75.36	76.2	6.3	2.6	2.5	0.123	1.392
Shawville QC	45.62	-76.47	167.6	4.8	2.4	2.5	-0.049	1.208
Sheenboro QC	45.97	-77.25	137.2	4.8	2.3	2.6	-0.267	1.042
South Mountain	44.97	-75.48	84.7	6.2	2.6	2.4	0.124	1.364
Wright QC	46.07	-76.05	141.7	4.4	2.6	2.5	0.092	1.347
Chats Falls	45.47	-76.23	93.9	6.1	2	2.3	-0.304	0.891
Chenaux	45.58	-76.68	84.1	5.8	2.2	2.5	-0.218	1.039
Chute St Phillipe QC	46.65	-75.25	251.5	3	2.9	2.5	0.43	1.709
Claybank	45.42	-76.4	106.7	5.6	2.2	2.4	-0.188	1.019
Crystal Falls	46.45	-79.87	227.1	3.5	2.5	2.4	0.115	1.357
French RC Dam	46.13	-80.01	198.1	4.7	2.5	2.4	0.077	1.294
Katrine Doe Lake	45.53	-79.4	305.1	5.5	3.1	2.3	0.784	1.968
Lake Traverse	45.95	-78.07	236.2	4	2.2	2.5	-0.365	0.925
Maniwaki UA QC	46.3	-76.01	188.6	3.7	2.4	2.5	-0.022	1.231
Notre Dame Du Laus QC	46.12	-75.63	213.4	4.1	2.6	2.3	0.316	1.508
Petawawa A	45.95	-77.32	130.1	4.1	2.2	2.6	-0.441	0.896
Petawawa Nat Forestry	45.98	-77.43	183	4.3	2.3	2.4	-0.115	1.109
Purdy	45.32	-77.72	490.7	4.6	2.3	2.4	-0.097	1.147
Trout Creek	46.03	-79.38	333.1	4.2	2.8	2.4	0.469	1.677
Wakefield QC	45.6	-75.9	152.4	5.1	2.7	2.5	0.234	1.513
Waltham QC	45.92	-76.92	112.5	4.1	2.5	2.6	-0.124	1.213
West Guilford	45.1	-78.68	327.2	4.1	3.3	2.5	0.825	2.104
Drummond	45.03	-76.25	145	6.1	2.3	2.6	-0.267	1.065
St Albert	45.29	-75.06	80	6	2.8	2.5	0.388	1.638
Ste Anne Du Lac QC	45.85	-75.33	262.1	3.1	2.8	2.4	0.387	1.616
Bancroft	45.05	-77.85	327	4.8	2.4	2.5	-0.081	1.198
Average of 51 stations				**4.8**	**2.6**	**2.5**	**0.144**	**1.394**

The compiled dataset (Table 4.3) was open in ArcGIS after the calculation of ETo, MA_1, and MA_2 with historical climate normals data (average of 1971–2010) for 51 meteorological stations, and a point feature map was created.

Almost half the weather station (Tables 4.2 and 4.3) results showed negative moisture availability (MA_1), meaning there was a moisture deficit during the months from April to September (Table 4.1, highlighted), even though for some stations, an increase in precipitation was also accompanied by a moisture deficit due to the rise in temperature resulting in higher evapotranspiration (ETo).

Figure 4.3 shows the spatial distribution of the 51 climate stations located throughout the study area, including in the rectangular region surrounding the study area extending into the neighboring province of Quebec. It also presents all the attributes of the Madawaska Weather Station (the geostatistical temperature layer map was created using 1971–2010 climate normals data).

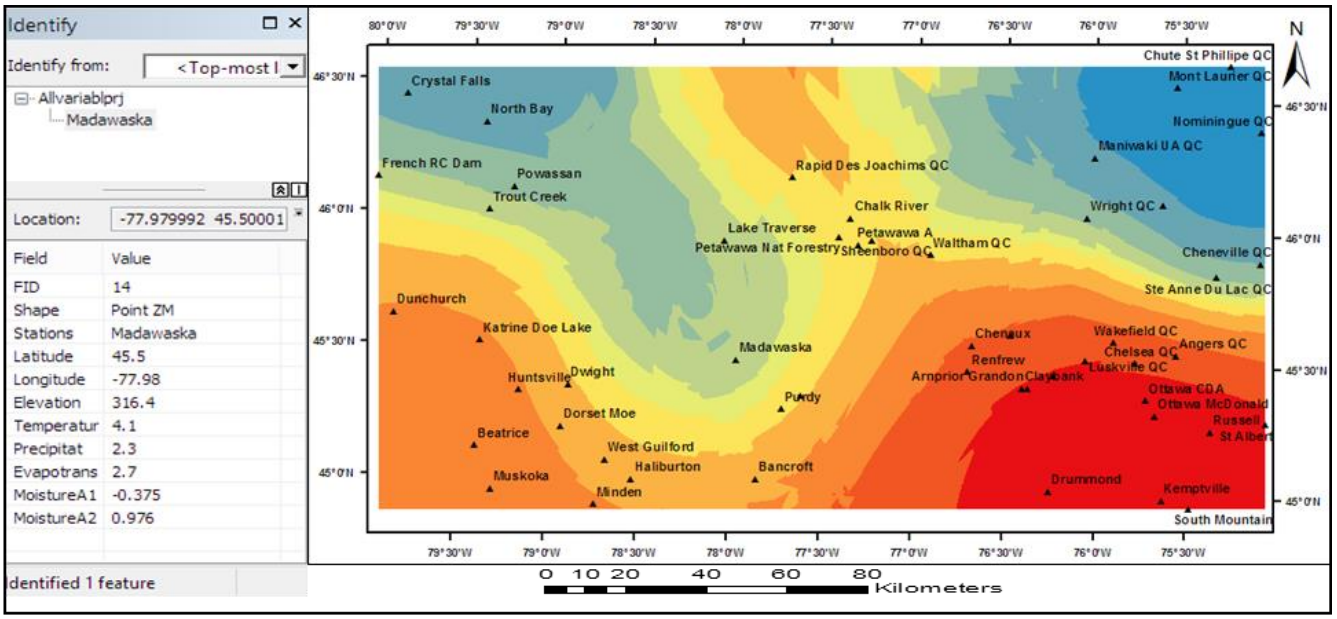


Figure 4. 3: Spatial distribution of the 51 climate stations located throughout the rectangular region of the study area (Note the point features of the geostatistical layer temperature map showing all attributes of the Madawaska station)

4.4.2 Forest disturbance and degradation trajectory maps 1972–2020 (LandTrendr product)

Long-term historical forest degradation trajectory maps were created for the periods from 1972–1983, 1984–2000, and 2001–2020 using the LandTrendr algorithm (details in Chapter 3). To ascertain the long-term degradation trend in the selected area, these three degradation maps were combined into one for the period from 1972–2020 using the image fusion technique (PCA) (Figure 4.4 (a-b)) (details in Chapter 3). The graduated color in Figure 4.4 indicates the low to high degradation areas during the period from 1972–

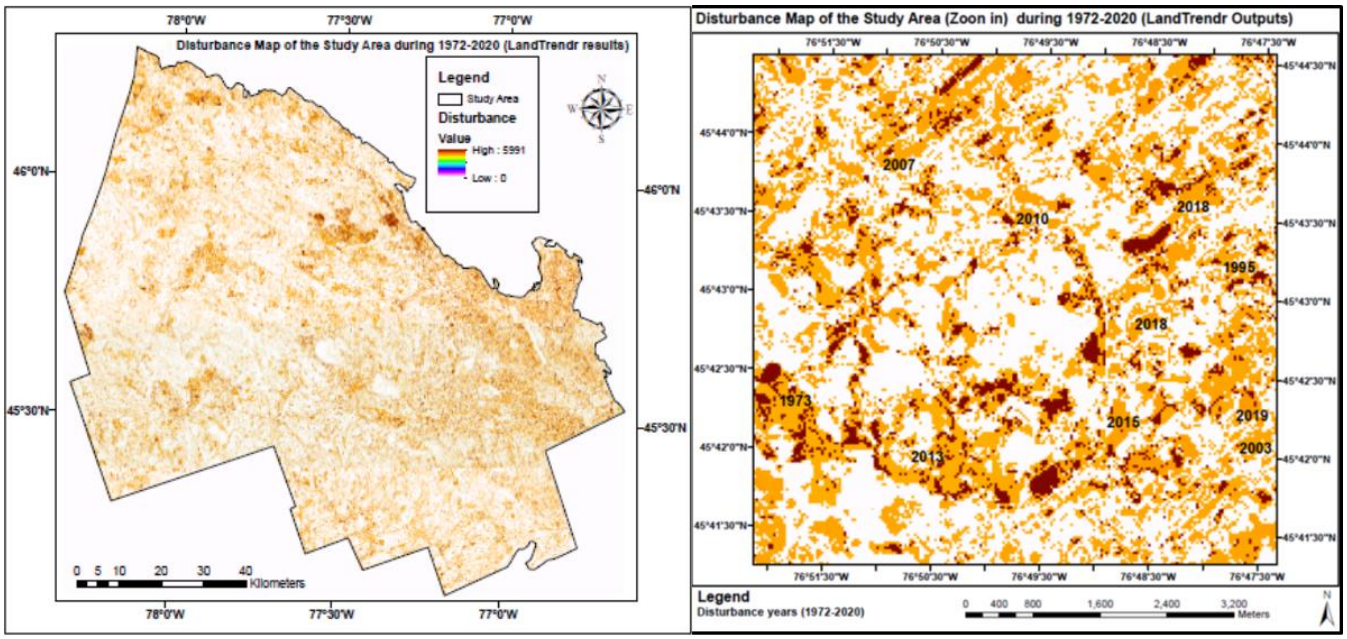


Figure 4. 4 (a-b): a) Disturbance/degradation map (1972–2020) of the study area; b) Large-scale disturbance/degradation map (1972–2020)

2020 (4a), and the large-scale map indicates different years of degradation (4b).

4.4.3 Historical (1971–2010) interpolated maps of climatology

ETo, MA_1, and MA_2 computed with historical climate normals data (average of 1971–2010) from 51 meteorological stations (point feature map) were spatially interpolated to generate prediction/gridded maps of Elev, T, P, ETo, MA_1, and MA_2 using the EBK interpolation procedure in ArcGIS (Figures 4.5 a-b, 4.6 c-d, and 4.7 e-f).

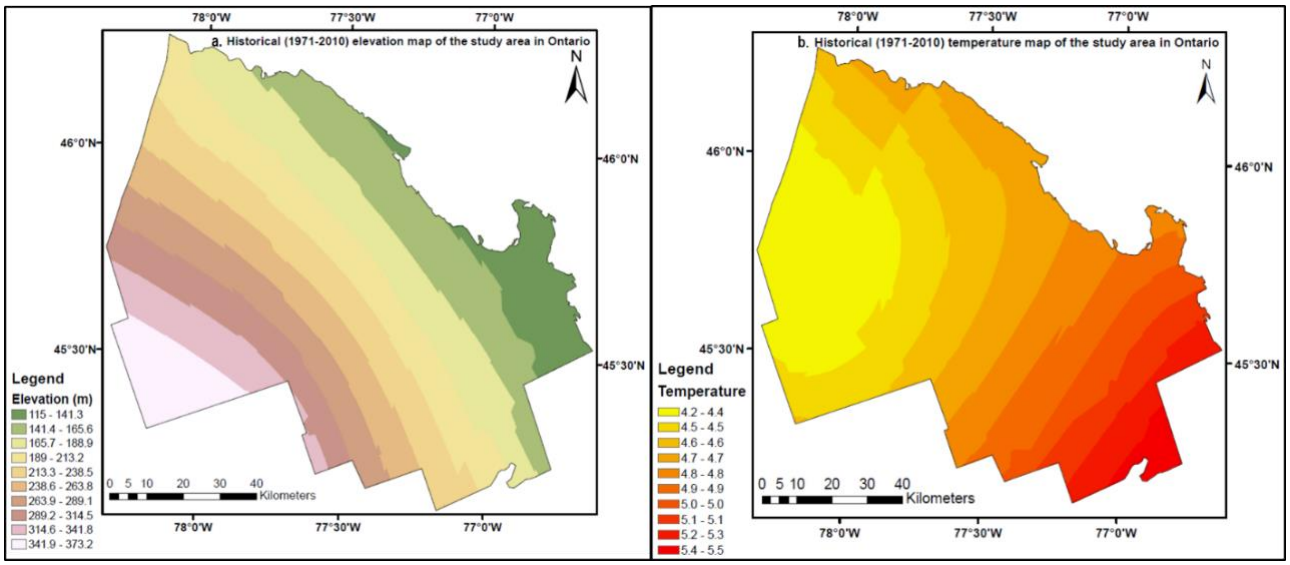


Figure 4. 5 (a-b): (a) Elevation map and (b) Spatial distribution/pattern of temperature maps (historical climate normals for 1971–2010)

Figure 4.5(a) shows the interpolated map of the elevation of the study area, which ranged from 115m to 373.2m. Higher elevations are found in the southwest side and lower elevations in the eastern side of the study area. The graduated color ramp of the interpolated map for the 40-year average temperature (Figure 4.5b) indicated lower temperatures in the western part (higher elevation area) and the highest temperatures in the southeastern parts of the study area (4.2^0C to $5.\ 5^0$C).

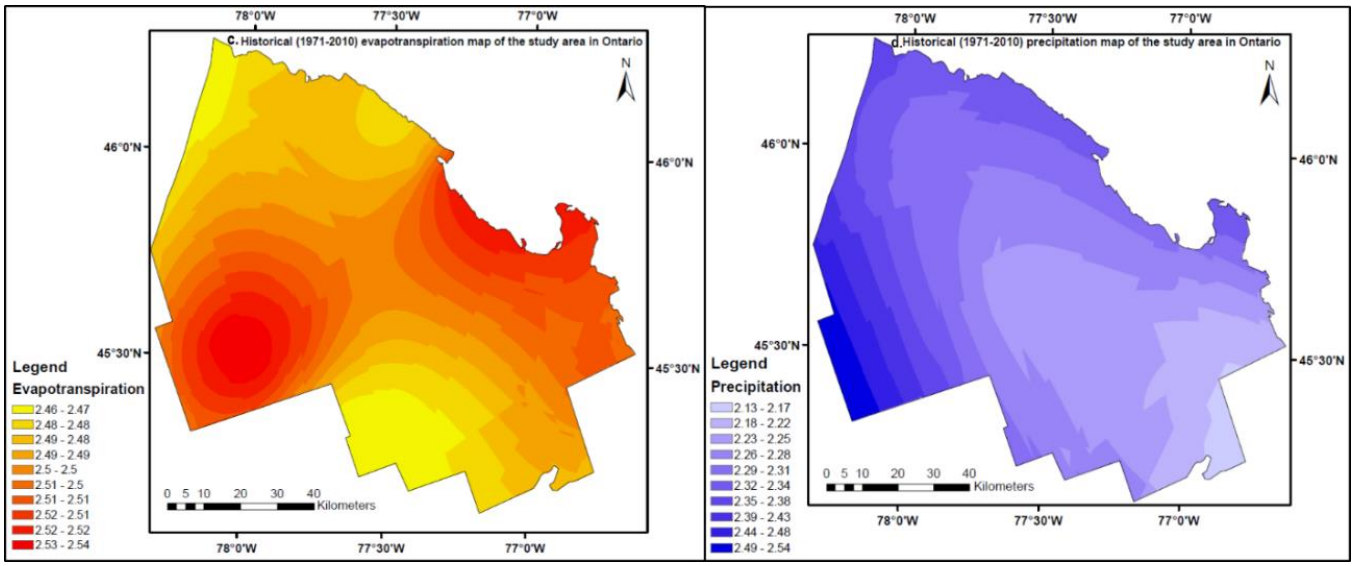


Figure 4. 6 (c–d): Spatial distribution of the evapotranspiration (ETo) (c) and precipitation (d) prediction maps for historical climate normals (1971–2010)

The graduated color ramp combination for the 40-year daily average ETo (2.46mm–2.54mm) interpolated map (4.6c) highlights the higher ETo in the southwestern and mid-eastern part, as well as moderate ETo in other parts of the study area. The graduated color ramp of light blue to dark blue for daily average precipitation (2.13mm–2.54mm) shows lower to higher precipitation for the interpolated map (4.6d) and 40-year daily average indicating lower precipitation in the southeastern parts, gradually increasing in the middle portions, and high precipitation in the western parts of the study area.

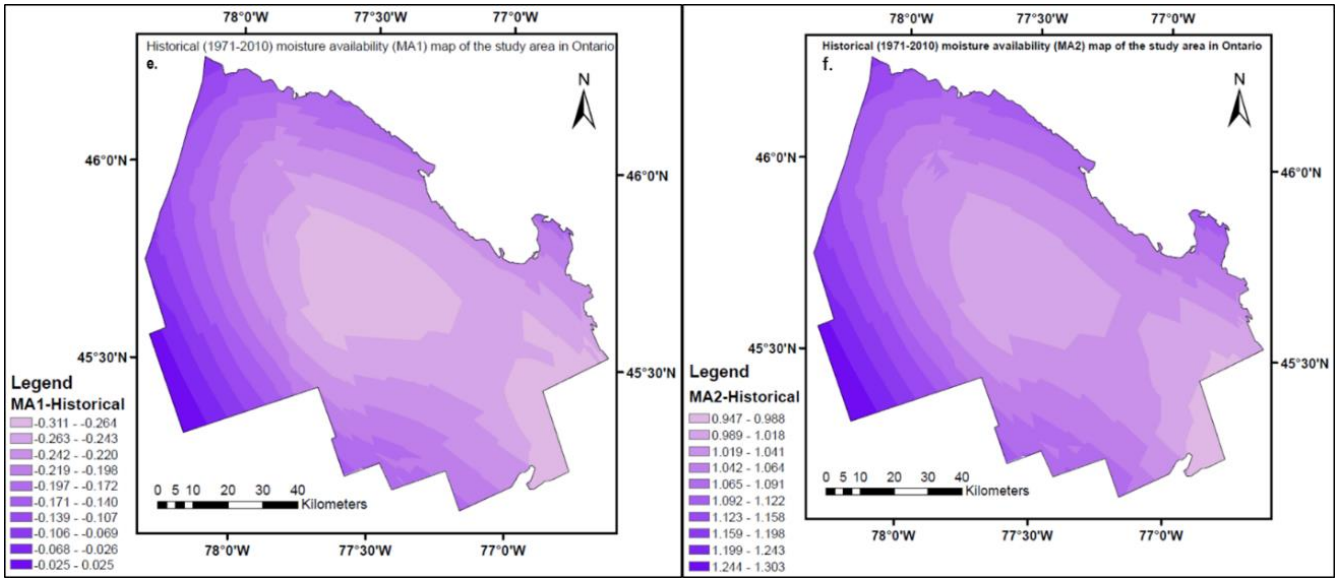


Figure 4. 7 (e–f): Spatial distribution of MA_1 (e) and MA_2 (f) interpolated maps (historical climate normals for 1971–2010)

The graduated color ramp of light blue to dark blue for daily average MA_1 (-0.311mm to 0.025mm) reveals the lower to higher MA_1 deficit of the interpolated map (4.7e) and 40-year daily average indicating a lower moisture deficit in the south and eastern parts, gradually increasing in the middle section, and a high deficit in the western parts of the study area. The graduated color ramp of MA_2 (4.7f) (0.917 to 1.303) also indicated a similar direction for MA_1.

4.4.4 Historical (2001–2020) remote sensing-derived maps (CFDI, NDVI, NPP)

Historical composite remote sensing maps (Figure 4.8 a-c) of CFDI, NDVI, and NPP values are a combination of 20 years (2001–2020) of map data. Overall CFDI involves seven FDIs (NDVI, NPP, NBR, EVI, TCG, TCB, and TCW) for each year in a 20-year period combined using PCA. NDVI and NPP are also the combination of 20 years of NDVI and NPP, respectively.

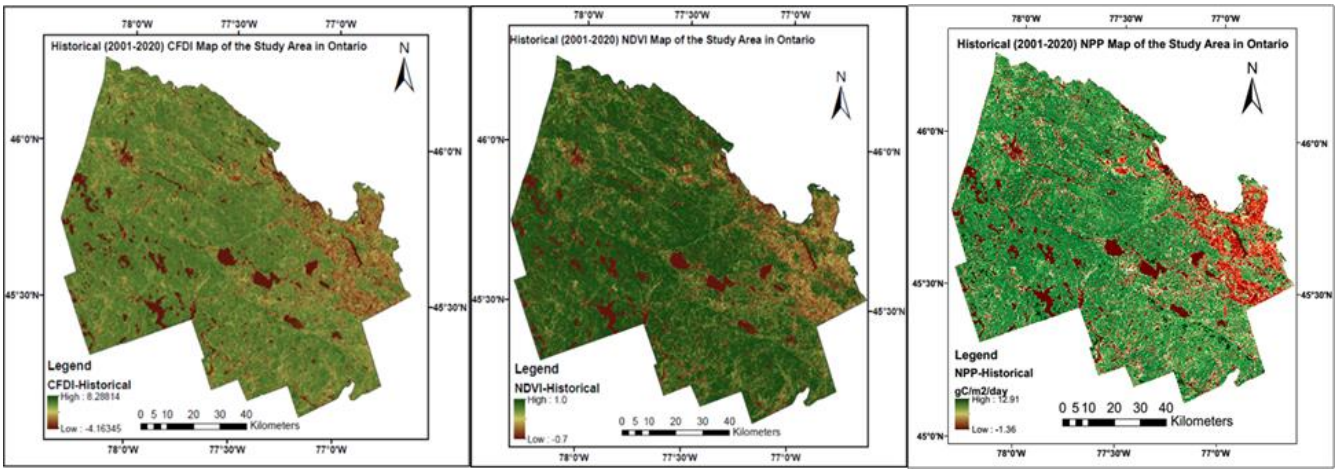


Figure 4. 8 (a-c): Historical (2001–2020) composite remote sensing maps of CFDI (a), NDVI (b), and NPP (c)

In Figure 4.8(a), there is no degradation in the dark green areas; the light green areas indicate medium/moderate degradation; and red areas indicate more degradation. For the historical NDVI map (Figure 4.8b), dark green indicates high NDVI (no degradation); light green represents medium NDVI (a little degradation); and red denotes a low NDVI (highly degraded areas). The graduated single-color ramp map (Figure 4.8c) of NPP employs deep green for no degradation; light green denotes medium/moderate degradation; and no green represents strong/more degradation.

4.4.5 Developing multiple linear regression models from site-specific pixel selection from fully overlapping maps (1972–2020)

The historical CFDI and prediction maps of climatology were fully overlapped pixel to pixel onto the forest disturbance and degradation map for site-specific pixel selection and analysis, and the pixel values were recorded for three different periods from 1972–1981, 1984–1998, and 2001–2020. These three

datasets were combined into one for the period from 1972–2020 to examine the potential statistical relationships between CFDI and climate variables. These are shown in Table 4.4.

Table 4. 4: Combined pixel values of forest disturbance/degradation for the study area (1972–2020)

Station Name	**Long**	**Lat**	**T**	**P**	**ETo**	**MA_1**	**MA_2**	**CFDI**
Petawawa A	-77.32	45.91	4.6	2.29	2.51	-0.238	1.048	1.112
Petawawa Nat Forestry	-77.66	45.96	4.5	2.28	2.48	-0.245	1.032	0.943
Chalk River	-77.38	46.05	4.6	2.32	2.48	-0.194	1.089	0.703
Rolphton	-77.94	46.16	4.5	2.32	2.48	-0.178	1.090	1.228
Lake Travers	-78.08	45.91	4.3	2.31	2.48	-0.208	1.057	1.447
Madawaska	-78.00	45.50	4.3	2.39	2.54	-0.139	1.134	-0.165
Gunters/South AP	-78.14	45.34	4.4	2.51	2.50	-0.004	1.273	-1.079
Combermere	-77.62	45.37	4.6	2.27	2.46	-0.211	1.060	1.180
Renfrew	-77.02	45.44	5.0	2.21	2.49	-0.258	1.017	2.009
Chenaux	-76.68	45.58	5.2	2.21	2.50	-0.259	1.010	1.696
Spotswood	-76.84	45.86	4.7	2.34	2.52	-0.192	1.101	0.431

To examine the relationship between the CFDI and climate variables, a Pearson correlation matrix (Table 4.5) was created based on the pixel values of the long-term (1972–2020) forest degradation trajectory map, with the CFDI derived from remote sensing and climate variables for the period from 1972–2020. For this study, an arbitrary criterion was established for the acceptance of the correlation coefficient (r) value as significant. Correlation coefficients greater than or equal to 0.75 ($r \geq 0.75$) were deemed significant. Variable values that are shaded and presented in bold denote that they are highly correlated.

Table 4. 5: Correlation matrix of CFDI and climate variables

	T	**P**	**ETo**	**MA_1**	**MA_2**	**CFDI**
T	1					
P	-0.66	1				
ETo	-0.06	0.36	1			
MA_1	-0.53	**0.97**	0.27	1		
MA_2	-0.52	**0.98**	0.31	**0.99**	1	
CFDI	0.57	**-0.95**	-0.46	**-0.91**	**-0.93**	1

Regression analysis was conducted based on the highly correlated dependent variable (CFDI) and independent climate variables (P, MA_1, and MA_2). The resulting equations are shown in Table 4.6.

Table 4. 6: Regression equation, projected models, and evaluation of model performance (1972–2020)

Indicators	Regression equations and predicted models	Validation	
		R^2	RMSE/SE
CFDI vs P	CFDI= – 10.067 P + 24.153	0.91	-
	CFDI ($\hat{Y}$) = 24.153 – 10.067 P	0.91	0.28
CFDI vs MA1	CFDI = – 9.8271x – 1.4512	0.81	-
	CFDI ($\hat{Y}$) = –1.249 – 10.929 MA1	0.82	0.39
CFDI vs MA2	y = – 9.8514x + 11.116	0.83	-
	CFDI ($\hat{Y}$) = 12.942 – 11.154 MA2	0.87	0.34
Predicted CFDI	CFDI ($\hat{Y}$) = 42.532–8.888 P +15.625 MA1–16.703 MA2	0.93	0.28

For example, a scatter plot, including regression lines and their computed coefficients of determination (R^2), for the dependent and independent variables, such as CFDI vs MA_1, were created (Figure 4.9). A model for the prediction of the CFDI as a function of MA_1 resulting from P-ETo is shown in Figure 4.9.

CFDI = –10.929 MA_1–1.249, where r = –0.91, R^2=0.82, RMSE=0.39

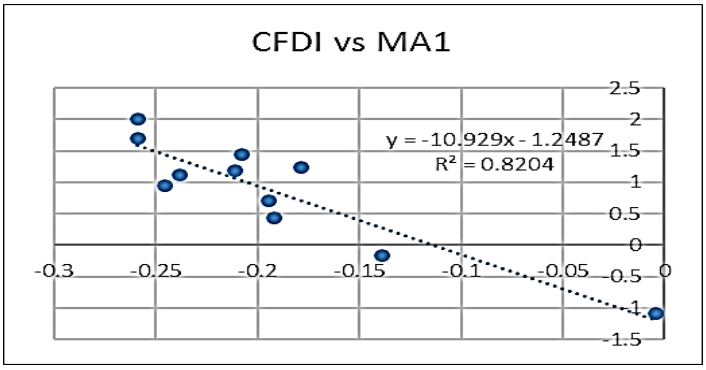


Figure 4. 9 Scatter plot of CFDI vs MA_1

In the multiple regression analysis, the best-fitted relationship included the CFDI with climate variables P, MA_1, and MA_2 (R^2=0.93 and RMSE=0.28). There was an expected high level of correlation among the climate variables denoting moisture availability (MA_1, MA_2) and precipitation over long periods, with the CFDI, but ETo and temperature were not individually correlated (Table 4.5). A very strong correlation between CFDI and moisture availability (MA_1=-0.91 and MA_2=-0.93) was found. When the temperature variable was included in the regression analysis (see the Materials and Methods section), the coefficients were either not significant or denoted relationships opposite to those with the MA variables. In all cases,

multiple regressions involving temperature and ETo led to weaker relationships and lower R^2 (as well as r) values than those based on moisture availability.

4.4.6 Trajectory analysis of forest disturbance and degradation maps 2001–2020 (LandTrendr product)

The results of 20 years of disturbance trajectory analysis (2001–2020) used in the LandTrendr algorithm to generate "disturbance maps" of the area at small and large scales are shown in Figure 4.10

Each block of color in Figure 4.10 (a–b) indicates the disturbance/degradation or change in landscape

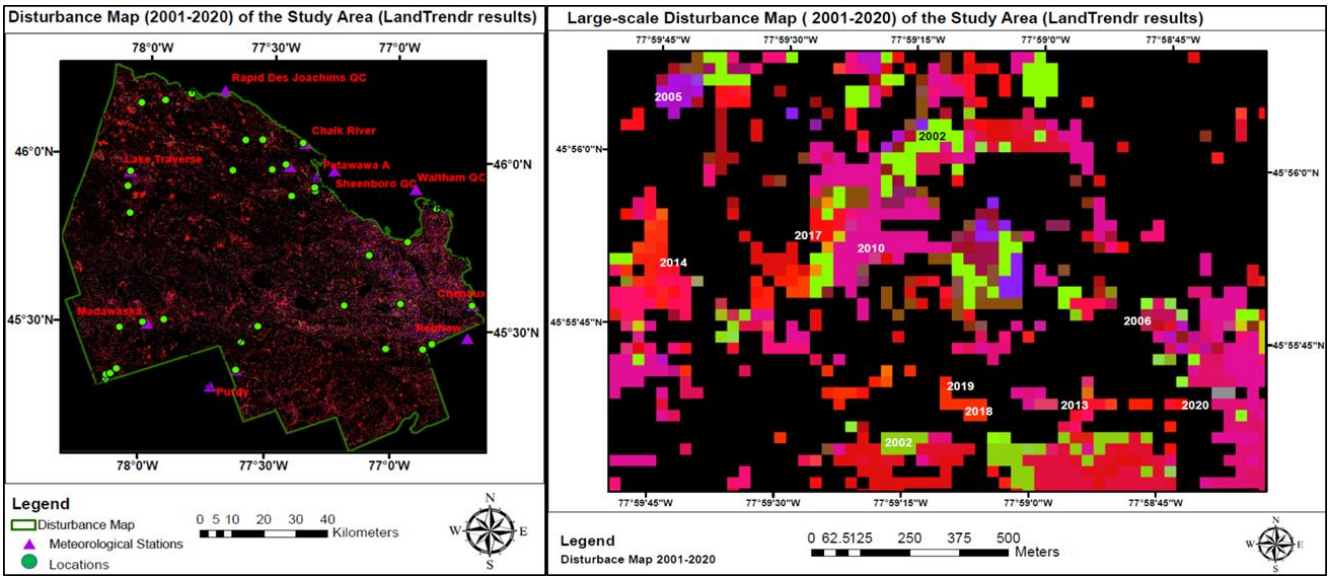


Figure 4. 10 (a-b): (a) Disturbance map (2001–2020) of the study area; (b) Large-scale disturbance map

during the period from 2001–2020, and the various colors denote different years of disturbance/degradation or change on the large-scale map (Figure 4.10b) above.

4.4.7 Developing multiple linear regression models from site-specific pixels from fully overlapped maps (2001–2020)

The CFDI for 2001–2020 is the most informative FDI, and a combination of seven FDIs (NDVI, NPP, NBR, EVI, TCB, TCG, and TCW) was also included in the analysis.

The historical (2001–2020) remote sensing (CFDI, NDVI, NPP) maps and historical (1971–2010) prediction maps of climatology were fully overlapped pixel-to-pixel to the forest disturbance/degradation

map (2001–2020) in order to conduct a site-specific pixel selection. The pixel values from these raster maps were recorded to examine the relationship between FDIs and climate variables (Table 4.7).

Table 4. 7: Summary of all fully overlapped map values for the study area (2001–2020)

Station Name	Long	Lat	T	P	ETo	MA_1	MA_2	NDVI	NPP	CFDI
Petawawa A	-77.33	45.92	4.6	2.29	2.51	-0.236	1.050	1.24	0.88	0.88
Petawawa Nat Forestry	-77.66	45.96	4.5	2.28	2.48	-0.245	1.032	0.54	0.86	0.65
Chalk River	-77.61	46.05	4.6	2.30	2.48	-0.213	1.065	0.40	0.68	0.49
Rolphton	-78.04	46.15	4.5	2.32	2.47	-0.168	1.096	1.01	0.58	0.73
Lake Travers	-78.07	45.95	4.3	2.31	2.48	-0.201	1.062	0.17	0.89	0.98
Madawaska	-78.00	45.50	4.3	2.39	2.54	-0.139	1.134	-0.36	-0.45	-0.46
Gunters/South AP	-78.12	45.35	4.4	2.50	2.50	-0.016	1.261	-0.69	-0.94	-1.34
Combermere	-77.54	45.50	4.6	2.25	2.48	-0.249	1.027	0.98	1.49	0.48
Renfrew	-77.02	45.44	5.0	2.21	2.49	-0.258	1.017	1.70	1.38	1.69
Chenaux	-77.19	45.57	4.8	2.24	2.49	-0.258	1.023	1.19	1.33	0.69
Spotswood	-76.95	45.76	4.8	2.28	2.51	-0.235	1.052	0.30	0.46	0.14

A Pearson correlation coefficients matrix (Table 4.8) was created based on all climate variables—including CFDI, NDVI, and NPP—for the period from 2001–2020 to examine the strength of the relationships between indicators derived from remote sensing and climate variables.

Table 4. 8: Correlation matrix of all variables (climate variables were measured from 1971–2010 climate normals data; CFDI, NDVI, and NPP variables were measured from 2001–2020 satellite data)

	T	P	ETo	MA_1	MA_2	NDVI	NPP	CFDI
T	1							
P	-0.67	1						
ETo	-0.09	0.40	1					
MA_1	-0.58	**0.98**	0.33	1				
MA_2	-0.54	**0.98**	0.38	**0.99**	1			
NDVI	**0.70**	**-0.85**	-0.43	**-0.78**	**-0.79**	1		
NPP	0.56	**-0.95**	-0.59	**-0.91**	**-0.93**	**0.87**	1	
CFDI	0.51	**-0.87**	-0.50	**-0.83**	**-0.86**	**0.86**	**0.88**	1

Regression analysis was conducted based on the highly correlated dependent (CFDI, NDVI, NPP) and independent (P, MA_1, and MA_2) variables to develop predictive MLR models using the historical satellite image-based map data and historical gridded prediction map of climate normals data (1971–2010), within

the disturbance/degradation area, identified by the LandTrendr tool. The regression equations fit yielded the regression line and R^2, which are shown in Table 4.9

4.4.8 Correlations between FDIs and climate variables

The relationship between CFDI/FDIs and climate variables (Tables 4.5 and 4.8), as previously discussed in the methodologies, is open to various interpretations depending on the variable pairs chosen. Their direct relevance to the study objectives also varies.

Correlations between different FDIs (Table 4.8) provide some indication of confidence in the different indicators; however, interpolations must be done carefully as these indicators are not necessarily independent in their derivation and correlation (or lack thereof) leads to the question of which one provides the most meaningful information for study goals. Additionally, owing to its stronger central place within the correlation matrix, CFDI was considered as the best degradation indicator for the interpretation.

The correlations between climate variables and degradation indicators (Tables 4.5 and 4.8) are likely to provide the most direct insights related to the objective of this study. In particular, they might furnish evidence of naturally occurring changes in the environmental landscape and also help distinguish the effects of temperature and precipitation.

The correlation coefficients between climate variables and FDIs were calculated in order to ascertain the climate variables most highly correlated to the CFDI in the study area; meaning that they would, of course, best represent the actual conditions (Table 4.9).

Table 4.9 shows the regression equations based on highly correlated FDIs, climate variables, and equation validations. The regression model's efficiency was evaluated on the basis of the smallest RMSE value and the highest coefficient of determination value, R^2.

Table 4. 9: Regression equations, predicted models, and evaluation of model performance (2001–2020)

Indicators	**Regression equations and predicted models**	**Validation**	
		R^2	**RMSE/SE**
CFDI vs P	y = – 8.6481x + 20.455	0.82	–
CFDI vs MA_1	y = – 9.8271x – 1.4512	0.81	–
CFDI vs MA_2	y = – 9.8514x + 11.116	0.83	–
Predicted CFDI	$CFDI(\hat{Y})$= 0.9506 NDVI – 0.1125	0.74	0.42
Predicted CFDI	$CFDI(\hat{Y})$= 0.9318NPP – 0.1578	0.77	0.40
Predicted NPP	$NPP(\hat{Y})$= 0.903 NDVI + 0.118	0.75	0.39
Predicted CFDI	$CFDI(\hat{Y})$=17.082–3.092P+2.002MA_1–8.398MA_2	0.83	0.39
Predicted NDVI	$NDVI(\hat{Y})$=29.206–14.654P+2.755MA_1+5.349 MA_2	0.80	0.38
Predicted NPP	$NPP(\hat{Y})$=38.90–10.464P+12.216 MA_1–10.848 MA_2	0.92	0.26

The paired statistics affect the variation in the rationale for correlation. First, the degree of correlation between the different FDIs (CFDI, NDVI, and NPP for the period from 2001–2020; Table 4.9) shows their robustness. In addition, the correlation between the climate variables (T, P, ETo, MA_1, and MA_2; Table 4.5) reveals the nature of climate forcing. Finally, the correlations (or their absence) between degradation indicator variables and climate variables may distinguish the natural/climatic and anthropogenic factors behind landscape change. The various relationships are bolded differently in Table 4.5 and Table 4.8 to help clarify their interpretation.

Correlations among CFDI, NDVI, NPP, T, P, ETo, MA_1, and MA_2 over time are summarized for the time series in Table 4.8. The CFDI shows a strong positive correlation with NDVI and NPP indicators. The correlation between CFDI, T, and ETo was weaker, but still significant, at $\alpha < 0.05$, and that between NDVI, NPP, and T was positive, although below the threshold (r or R^2>0.75) of significance. In accordance with expectations, the climatic variables correlate inversely. However, no significant correlations were observed between some climatic variables (T, ETo) and degradation indicators.

When looking at the correlation (or lack thereof) between climate and degradation indicator variables, the role of seasonality in the often hot/humid and semi-arid climate of the study area becomes clear. Of the

climate variables, CFDI is the most closely linked to the forest vegetation response, and it is not strongly correlated with temperature and ETo, but it is strongly correlated with other variables (P, MA_1, and MA_2). Finally, two historical (1972–2020 and 2001–2020) CFDI gridded maps were created (Figure 4.11 a-b) using the following predicted MLR models, calculated by the raster calculator tool in ArcGIS (using historical gridded maps of P, MA_1, and MA_2, shown in Figures 4.6d and 4.7e-f).

$$\text{Predicted CFDI } (\hat{Y}) = 42.532 - 8.888P + 15.625MA_1 - 16.703MA_2 \quad (4.5)$$

$$\text{Predicted CFDI } (\hat{Y}) = 17.082 - 3.092P + 2.002MA_1 - 8.398MA_2 \quad (4.6)$$

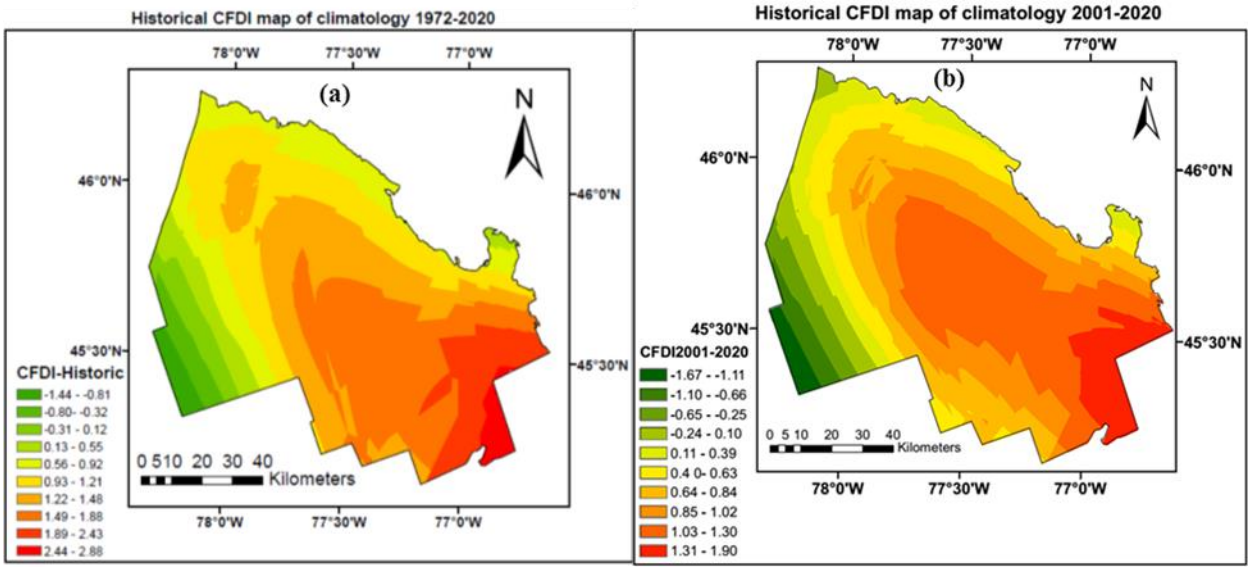


Figure 4. 11 (a-b): a) Historical CFDI gridded maps of climatology created based on the prediction models of (a) 1972–2020 (Eq 4.5) and (b) 2001–2020 (Eq 4.6)

A high level of degradation is presented in red; green denotes very low or no degradation; and shades in between red and green indicate a low level of degradation.

4.5 Discussion

The objective of this study was the development of an empirical model to assess the effects of climate change on forest cover in an ecologically important study area in Ontario by integrating field data, remote sensing, and climate projection to predict future forest degradation.

Although there have been numerous studies on the impact of climate change on vegetation cover change (Sun et al., 2020; Ruiz-Benito et al., 2020) and its response to global warming (Lawal et al., 2019; Ma et al., 2020), as well as the prediction of future LULC (Schirpke et al., 2017; Yang et al., 2019), there has been little exploration of identifying vulnerable areas by quantifying the effects of climate change on vegetation cover in the forest ecosystem. While this is not the first study to examine the impact of climate change on vegetation/forest cover, it recommends a new prediction MLR model that provides a forest vegetation cover degradation map under the projected climate scenario of 2050.

The findings for 1972–2020 indicated climate variables P, MA_1, and MA_2 had the strongest correlation (-0.95, -0.91, and -0.93, respectively) with CFDI, which enabled the computing of an MLR model with a high coefficient of determination, R^2 (0.93), and low RMSE (0.28). Additionally, findings for 2001–2020 showed climate variables P, MA_1, and MA_2 had the most significant correlation with CFDI, NDVI, and NPP. This strong relationship allowed for the establishment of future models with a high coefficient of determination, R^2 (0.83, 0.80, 0.92, respectively), and low RMSE (0.39, 0.38, 0.26, respectively).

Finally, projected MLR models were established based on the highly correlated FDIs (CFDI, NDVI, and NPP) and climate variables (P, MA_1, and MA_2).

Trends in temperatures in and close to the study area reveal a rise in daily average temperature, which was 4.9^0C, 5.1^0C, and 5.2^0C for 1971–2000, 1981–2010, and 2011–2020, respectively. Additionally, the Ontario study area forests displayed evidence of moisture stress, particularly in the summer months (June–August) from 1971–2010. The results furnished evidence that forest cover will likely experience losses and large fluctuations in the future, putting the area's ecosystem at risk. Therefore, this method of monitoring forest degradation through the integration of remote sensing tools could assist and certainly improve climate change risk management, in particular by providing data on climate impacts on forest/vegetation.

As a result of a great deal of effort, daily grids of mean temperature, precipitation, ETo, as well as MA (MA_1 and MA_2) at 30m spatial resolution were produced for the study area for 1971–2010. This dataset is a complete, multi-decadal, spatiotemporally comprehensive database suitable for regional climate analysis, risk assessment, ecosystem modelling, as well as planning and management purposes. Its production was achieved with a much higher station density, greater spatiotemporal resolution, and longer data record than was possible in many previous studies with gridded databases (McKenney et al., 2006). The gridded and interpolated climate maps reveal the weather patterns that reflect the seasonal atmospheric processes in the region under investigation. However, both the chosen spatial interpolation (e.g., algorithm, parameterization) and quality of input data affect the accuracy of our climatological grids.

4.6 Conclusions

Since climate change will likely be a driving force in the change in future forest cover in a wide variety of ecosystems, a range of remote sensing and climate variables was applied to assess the potential of forest degradation measurement through this research. In terms of environmental planning and management, this research bridges current knowledge related to the spatial distribution, temporal trends (over the short and long term), and patterns in the climate over the last few decades.

The best practice taken with any data modelling approach is to judge it by its efficacy, and the results of testing EBK models have shown that this method is a good approach for both real and simulated data. EBK-obtained prediction intervals have good coverage probabilities in instances when there are changes in the data variation that are rapid and also dissimilar in different parts of the data. It is common for EBK to surpass all other predictors in terms of its performance, especially with data complexity and when a number of typical geostatistical data modeling assumptions are violated.

This research analyzed the possible relationship between climate variables and FDIs in order to provide forest degradation prediction maps for the study area. As the results of losses and large fluctuations in

vegetation in the future pose risks to the ecosystem in the area, this method of monitoring forest degradation through the integration of tools could assist and certainly improve climate change risk management and planning, in particular by providing data on climate impacts.

References

Addis, H.K., Klik, A., and Strohmeier, S. 2016. Performance of frequently used interpolation methods to predict spatial distribution of selected soil properties in an agricultural watershed in Ethiopia. Appl. Eng. Agric., 32(5), pp.617–626.

Ajjur, S.B., and Al-Ghamdi, S.G. 2021. Evapotranspiration and water availability response to climate change in the Middle East and North Africa. Climatic Change, 166(3), pp.1–18.

Ali, G., Sajjad, M., Kanwal, S., Xiao, T., Khalid, S., Shoaib, F., and Gul, H.N. 2021. Spatial–temporal characterization of rainfall in Pakistan during the past half-century (1961–2020). Scientific Reports, 11(1), pp.1–15.

Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., et al. 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. For. Ecol. Manag., 259, pp.660–684.

Allen, R.G., Pereira, L.S., Raes, D., and Smith, M. 1998. Crop. Evapotranspiration: FAO Irrigation and Drainage Paper 56. FAO: Rome, Italy, pp. 20.

Anderegg, W.R.L., Hicke, J.A., Fisher, R.A., and Allen, C.D. 2015. Tree mortality from drought, insects, and their interactions in a changing climate. N. Phytol., 208, pp.674–683.

Arguez, A., and Vose, R.S. 2011. The definition of the standard WMO climate normal: The key to deriving alternative climate normals. Bulletin of the American Meteorological Society, 92(6), pp.699 –704.

Barnett, T.P., Adam, J.C., and Lettenmaier, D.P. 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. Nature, 438, pp.303–309.

Berner, L.T., Law, B.E., and Hudiburg, T.W. 2017. Water availability limits tree productivity, carbon stocks, and carbon residence time in mature forests across the western US. Biogeosciences, 14, pp.365–378.

Bolivar, E. 2020. Representing past and future hydro-climatic variability over multi-decadal periods in poorly-gauged regions: The case of Ecuador. Hydrology. Université Paul Sabatier - Toulouse III, English. NNT: 2020TOU30119. tel-03124408.

Braun, M., Klaas, T., Vieira, E., and Eng, P. 2015. Manitoba-Minnesota transmission project historic and future climate study. Manitoba Hydro.

Campbell, C., Campbell, I.D., and Hogg, E.H. 1994. Lake area variability across a climatic and vegetational transect in southeastern Alberta, Canada. Geographie Physique et Quaternaire, 48, pp.207–212.

Chen, F.-W., and Liu, C.-W. 2012. Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. Paddy Water Environ., 10, pp.209–222.

Chilès, J.P., and Delfiner, P. 1999. Geostatistics: Modeling spatial uncertainty. John Wiley & Sons, Inc., pp.695.

Cui, X., Gibbes, C., Southworth, J., Waylen, P., Cui, X., Gibbes, C., Southworth, J., and Waylen, P. 2013. Using remote sensing to quantify vegetation change and ecological resilience in a semi-arid system. Land, 2(2), pp.108–130.

De Keersmaecker, W., Lhermitte, S., Tits, L., Honnay, O., Somers, B., Coppin, P. 2015. A model quantifying global vegetation resistance and resilience to short-term climate anomalies and their relationship with vegetation cover. Glob. Ecol. Biogeogr., 24, pp.539–548.

Eichelmann, E., et al. 2018. The effect of land cover type and structure on evapotranspiration from agricultural and wetland sites in the Sacramento–San Joaquin River Delta. California Agricultural and Forest Meteorology, 256–257, pp.179–195.

Environment Canada. 1986a. Temperature and degree days. Climatic Atlas—Canada, Map Series 1, Canadian Government Publishing Centre, pp.12.

Environment Canada. 1986b. Precipitation. Climatic Atlas—Canada, Map Series 2, Canadian Government Publishing Centre, pp.52.

ESRI. 2019a. ArcGIS geostatistical analyst. https://www.esri.com/en-us/arcgis/products/geostatistical-analyst/overview

Eskandari Dameneh, H., Gholami, H., Telfer, M.W., Comino, J. R., Collins, A. L., and Jansen, J. D. 2021. Desertification of Iran in the early twenty-first century: Assessment using climate and vegetation indices. Scientific Reports, 11(1), 1–18.

Esri (Environmental Systems Research Institute). 2019. ArcGIS Pro Help. Redlands, CA, https://pro.arcgis.com/en/pro-app/help/main/welcome-to-the-arcgis-pro-app-help.htm

FAO. 1978–81. Report on the agro-ecological zones project. World Soil Resources Report 48, Food and Agriculture Organization of the United Nations, Rome, Italy.

FAO. 2012. The ETo Calculator Reference Manual Version 3.2. Land and Water Division. Food and Agriculture Organization of the United Nations, Rome, Italy.

Fatiha, B., Abdelkader, A., Latifa, H., Mohamed, E. 2013. Spatio temporal analysis of vegetation by vegetation indices from multi-dates satellite images: Application to a semi-arid area in Algeria. Energy Procedia, 36, pp.667–675.

Feng, Y., Cui, N., Zhao, L., Gong, D., and Zhang, K. 2017. Spatiotemporal variation of reference evapotranspiration during 1954–2013 in Southwest China. Quaternary International, 441, pp.129–139.

Feng, Y., Jia, Y., Cui, N.B., Zhao, L., Li, C., Gong, D.Z. 2017. Calibration of Hargreaves model for reference evapotranspiration estimation in Sichuan basin of southwest China. Agric. Water Manag., 181, pp.1–9.

Fensholt, R., Langanke, T., Rasmussen, K., Reenberg, A., Prince, S.D., Tucker, C., Scholes, R.J., Le, Q.B., Bondeau, A., Eastman, R., and Epstein, H. 2012. Greenness in semi-arid areas across the globe 1981–2007—an Earth Observing Satellite based analysis of trends and drivers. Remote Sensing of Environment, 121, pp.144–158.

Fu, B.J., Li, S.G., Yu, X.B., Yang, P., Yu, G.R., Feng, R.G., and Zhuang, X.L. 2010. Chinese ecosystem research network: Progress and perspectives. Ecol. Complex., 7, pp.225–233.

Gong, L., Xu, C-y., Chen, D., Halldin, S., and Chen, Y.D. 2006. Sensitivity of the Penman–Monteith reference evapotranspiration to key climatic variables in the Changjiang (Yangtze River) basin. J. Hydrol., 329, pp.620–629.

Goulden, M.L., and Bales, R.C. 2014. Mountain runoff vulnerability to increased evapotranspiration with vegetation expansion. Proc. Natl. Acad. Sci. U.S.A., 111, pp.14071–14075.

Goulden, M.L., and Bales, R.C. 2019. California forest die-off linked to multi-year deep soil drying in 2012–2015 drought. Nat. Geosci., 12, pp.632–637.

Goulden, M.L., Anderson, R.G., Bales, R.C., Kelly, A.E., Meadows, M., and Winston, G.C. 2012. Evapotranspiration along an elevation gradient in California's Sierra Nevada. J. Geophys. Res. Biogeosci., 117, pp.1–13.

Guenni, L., and Hutchinson, M.F. 1998. Spatial interpolation of the parameters of a rainfall model from ground-based data. J. Hydrol., 212, pp.335–347.

Gupta, A., Kamble, T., and Machiwal, D. 2017. Comparison of ordinary and Bayesian kriging techniques in depicting rainfall variability in arid and semi-arid regions of north-west India. Environ. Earth Sci., 76(512).

Hargreaves, G.H., and Samani, Z.A. 1982. Estimating potential evapotranspiration. J. Irrig. Drain. Eng 108, pp.225–230.

Hogg, E.H. 1994. Climate and the southern limit of the western Canadian boreal forest. Can. J. For. Res., 24, pp.1835–1845.

Hogg, E.H. 1997. Temporal scaling of moisture and the forest-grassland boundary in western Canada. Agricultural and Forest Meteorology 84, pp.115–122.

Hogg, E.H., Barr, A.G., and Black, T.A. 2013. A simple soil moisture index for representing multi-year drought impacts on aspen productivity in the western Canadian interior. Agricultural and Forest Meteorology, 178–179, pp.173–182.

Hunter, C., Moore, R.D., and McKendry, I. 2020. Evaluation of the North American Regional Reanalysis (NARR) precipitation fields in a topographically complex domain. Hydrol. Sci. J., 65(5), 786–799.

Hutchinson, M.F., McKenney, D.W., Lawrence, K., Pedlar, J.H., Hopkinson, R.F., Milewska, E., and Papadopol, P. 2009. Development and testing of Canada-wide interpolated spatial models of daily minimum–maximum temperature and precipitation for 1961–2003. Journal of Applied Meteorology and Climatology, 48(4), pp.725–741.

Irmak, S., Payero, J.O., Martin, D.L., Irmak, A., and Howell, T.A. 2006. Sensitivity analyses and sensitivity coefficients of standardized daily ASCE-Penman-Monteith equation. Journal of Irrigation and Drainage Engineering, 132(6), pp.564–578.

Islam, A.R.M.T., Shen, S., Yang, S., Hu, Z., and Chu, R. 2019. Assessing recent impacts of climate change on design water requirement of Boro rice season in Bangladesh. Theor. Appl. Clim. 138(1–2), pp.97–113.

Jerin, J.N., Islam, H.M., Islam, A.R.M., Shahid, S., Hu, Z., Badhan, M.A., Chu, R., and Elbeltagi, A. 2021. Spatiotemporal trends in reference evapotranspiration and its driving factors in Bangladesh. Theoretical and Applied Climatology, 144(1), pp.793–808.

Jiapaer, G., Liang, S.L., Yi, Q.X., and Liu, J.P., 2015. Vegetation dynamics and responses to recent climate change in Xinjiang using leaf area index as an indicator. Ecol. Indic., 58, pp.64–76.

Kitanidis, P.K. 1997. Introduction to geostatistics: Applications in hydrogeology. New York: Cambridge University Press.

Krakauer, N., Lakhankar, T., Anadón, J., Krakauer, N.Y., Lakhankar, T., and Anadón, J.D. 2017. Mapping and attributing normalized difference vegetation index trends for Nepal. Remote Sens., 9, pp.986.

Krivoruchko, K. 2012a. Empirical Bayesian kriging. ArcUser Fall 2012.

Krivoruchko, K., and Fraczek, W. 2015. Interpolation of data collected along lines.

Lai, Y., Dzombak, D.A., Lai, Y., and Dzombak, D.A. 2019. Use of historical data to assess regional climate change. J. Clim., 32(14), pp.4299–4320.

Lawal, Sh., Lennard, Ch., and Hewitson, B., 2019. Response of southern African vegetation to climate change at 1.5 and 2.0° global warming above the pre-industrial level. Climate Services Journal, 100134.

Li, J., and Heap, A.D. 2008. A review of spatial interpolation methods for environmental scientists. Geoscience Australia, pp. 137.

Li, J., and Heap, A.D. 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. Ecological Informatics, 6, pp.228–241.

Li, J., Heap, A.D., Potter, A., and Daniell, J.J. 2011. Application of machine learning methods to spatial interpolation of environmental variables. Environmental Modelling & Software, 26, pp.1647–1659.

Ma, Q., Zhu, L., Wang, J., Ju, J., Wang, Y., Lu, X., Kasper, Th., and Haberzettl, T. 2020. Late Holocene vegetation responses to climate change and human impact on the central Tibetan Plateau. Journal of Science of the Total Environment, 708, pp.135370.

Marchildon, G.P. (Ed.). 2009. A dry oasis: institutional adaptation to climate on the Canadian plains (Part-II). Regina, Saskatchewan; CPRC Press, University of Regina.

McKenney, D.W., Hutchinson, M.F., Papadopol, P., Lawrence, K., Pedlar, J., Campbell, K., Milewska, E., Hopkinson, R.F., Price, D., and Owen, T. 2011. Customized spatial climate models for North America. Bulletin of the American Meteorological Society, 92(12), pp.161–122.

McKenney, D.W., Pedlar, J.H., Papadopol, P., Hutchinson, M.F. 2006. The development of 1901–2000 historical monthly climate models for Canada and the United States. Agric. Forest Meteorol., 138, pp.69–81.

McKenzie, D., Gedalof, Z., Peterson, D.L., and Mote, P. 2004. Climatic change, wildfire, and conservation. Conserv. Biol., 18, pp.890–902.

Mirzaei, R., and Sakizadeh, M. 2016. Comparison of interpolation methods for the estimation of groundwater contamination in Andimeshk-Shush Plain, Southwest of Iran. Environmental Science and Pollution Research, 23, pp.2758–2769.

Montandon, L.M., and Small, E.E. 2008. The impact of soil reflectance on the quantification of the green vegetation fraction from NDVI. Remote Sens. Environ, 112, pp.1835–1845.

Monteith, J.L., and Unsworth, M.H. 2008. Principles of environmental physics. 3rd ed. Amsterdam: Academic Press.

Naoum, S., and Tsanis, I.K. 2003. Temporal and spatial variation of annual rainfall on the island of Crete, Greece. Hydrol. Process, 17, pp.1899–1922.

Natural Resources Canada. 2011a. The atlas of Canada.

Natural Resources Canada. 2011b. Long term mean climate grids for Canada and United States.

North, M.P., Stephens, S.L., Collins, B.M., Agee, J.K., Aplet, G., Franklin, J.F., et al. 2015a. Reform forest fire management. Science, 349, pp.1280–1281.

Olmos-Trujillo, E., González-Trinidad, J., Júnez-Ferreira, H., Pacheco-Guerrero, A., Bautista-Capetillo, C., Avila-Sandoval, C., and Galván-Tejada, E. 2020. Spatio-temporal response of vegetation indices to rainfall and temperature in a semiarid region. Sustainability, 12(5), pp.1939.

Piedallu, C., Chéret, V., Denux, J.P., Perez, V., Azcona, J.S., Seynave, I., and Gégout, J.C. 2019. Soil and climate differently impact NDVI patterns according to the season and the stand type. Sci. Total Environ., 651, pp.2874–2885.

Price, D.T., McKenney, D.W., Nalder, I.A., Hutchinson, M.F., and Kesteven, J.L. 2000. A comparison of two statistical methods for spatial interpolation of Canadian monthly climate data. Agric. For. Meteor., 101, pp.81–94.

Prince, S. D., Becker-Reshef, I., and Rishmawi, K. 2009. Detection and mapping of long-term land degradation using local net production scaling: Application to Zimbabwe. Remote Sens. Environ., 113(5), pp.1046–1057.

Ren, S.L., Li, Y.T., and Peichl, M., 2020. Diverse effects of climate at different times on grassland phenology in mid-latitude of the Northern Hemisphere. Ecol. Indic., 113, pp.106260.

Roche, J. W., Ma, Q., Rungee, J., and Bales, R.C. 2020. Evapotranspiration mapping for forest management in California's Sierra Nevada. Frontiers in Forests and Global Change, 3, pp.69.

Ruiz-Benito, P., Vacchiano, G., Lines, E., Reyer, C., Morin, X., Hartig, F., Makela, A., Yousefpour, R., Chaves, J., Palacios-Orueta, A., Benito-Garzon, M., Molino, C., Camarero, J., Jump, A., Kattge, J., Lehtonen, A., Ibrom, A., and Zavala, M. 2020. Available and missing data to model impact of climate change on European forests. Ecological Modelling Journal, 416, pp.1–15.

Salam, R., and Islam, A.R.M.T. 2020. Potential of RT, Bagging and RS ensemble learning algorithms for reference evapotranspiration prediction using climatic data-limited humid region in Bangladesh. J. Hydrol., 590, pp.125241.

Schirpke, U., Kohler, M., Leitinger, G., Fontana, V., Tasser, E., and Tappeiner, U. 2017. Future impacts of changing land-use and climate on ecosystem services of mountain grassland and their resilience. Ecosystem Service Journal, 26, pp.79–94.

Shi, H.Y., Li, T.J., and Wang, G.Q. 2017. Temporal and spatial variations of potential evaporation and the driving mechanism over Tibet during 1961–2001. Hydrol. Sci. J., 62(9), pp.1469–1482.

Sun, P., Yu, Z., Liu, S., Wei, X., Wang, J., Zegre, N., and Liu, N. 2012. Climate change, growing season water deficit and vegetation activity along the north-south transect of eastern China from 1982 through 2006. Hydrol. Earth Syst. Sci., 16(10), pp.3835–3850.

Sun, Y., Shan, M., Pei, X., Zhang, X., and Yang, Y. 2020. Assessment of the impacts of climate change and human activities on vegetation cover change in the Haihe River basin, China. Journal of Physics and Chemistry of the Earth, Parts A/B/C, pp.102834.

Tabari, H., and Talaee, P.H. 2014. Sensitivity of evapotranspiration to climatic change in different climates. Glob Planetary Change, 115, pp.16–23.

Tague, C.L., Moritz, M.A., and Hanan, E. 2018. The changing water cycle: The eco-hydrologic impacts of forest density reduction in Mediterranean (seasonally dry) regions. WIRES Water 6:e1350.

Tang, B., Tong, L., and Kang, S.Z. 2013. Effects of spatial station density and interpolation methods on accuracy of reference crop evapotranspiration. Trans. Chinese Soc. Agric. Eng., 29(13), pp.60e66 (in Chinese).

Trajkovic, S., Kolakovic, S. 2010. Comparison of simplified pan-based equations for estimating reference evapotranspiration. J. Irrig. Drain. Eng., 136, pp.137–140.

Tveito, O.E., Wegehenkel, M., van der Wel, F., and Dobesch, H. 2008. COST Action 719: The use of geographic information systems in climatology and meteorology. EUR-OP, Luxembourg.

Valiantzas, J.D. 2013. Simplified forms for the standardized FAO-56 Penman-Monteith reference evapotranspiration using limited weather data. J. Hydrol., 505, pp.13–23.

Vermote, E., Justice, C., Claverie, M., and Franch, B. 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. Remote Sens. Environ., 185, pp.46–56.

World Meteorological Organization (WMO). 1989. Calculation of monthly and annual 30-years standard normals. World Climate Data and Monitoring Programme Series, WCDP 10, WMOTD 341.

World Meteorological Organization (WMO). 2007. The role of climatological normals in a changing climate. World Climate Data and Monitoring Programme Series, WCDMP-No 61 WMO-TD, No. 1377.

Yang, Q., Zhang, H., Peng, W., Lan, Y., Luo, S., Shao, J., Chen, D., and Wang, G. 2019. Assessing climate impact on forest cover in areas undergoing substantial land cover change using Landsat imagery. Science of the Total Environment Journal, 659, pp.732–745.

Yao, N., Li, Y., Lei, T.J., and Peng, L.L. 2018. Drought evolution, severity and trends in mainland China over 1961–2013. Sci. Total Environ., 616–617, pp.73–89.

Zhang, P.P., Cai, Y.P., Yang, W., Yi, Y.J., Yang, Z.F., and Fu, Q. 2019. Multiple spatio-temporal patterns of vegetation coverage and its relationship with climatic factors in a large dam-reservoir-river system. Ecol. Eng., 138, pp.188–199.

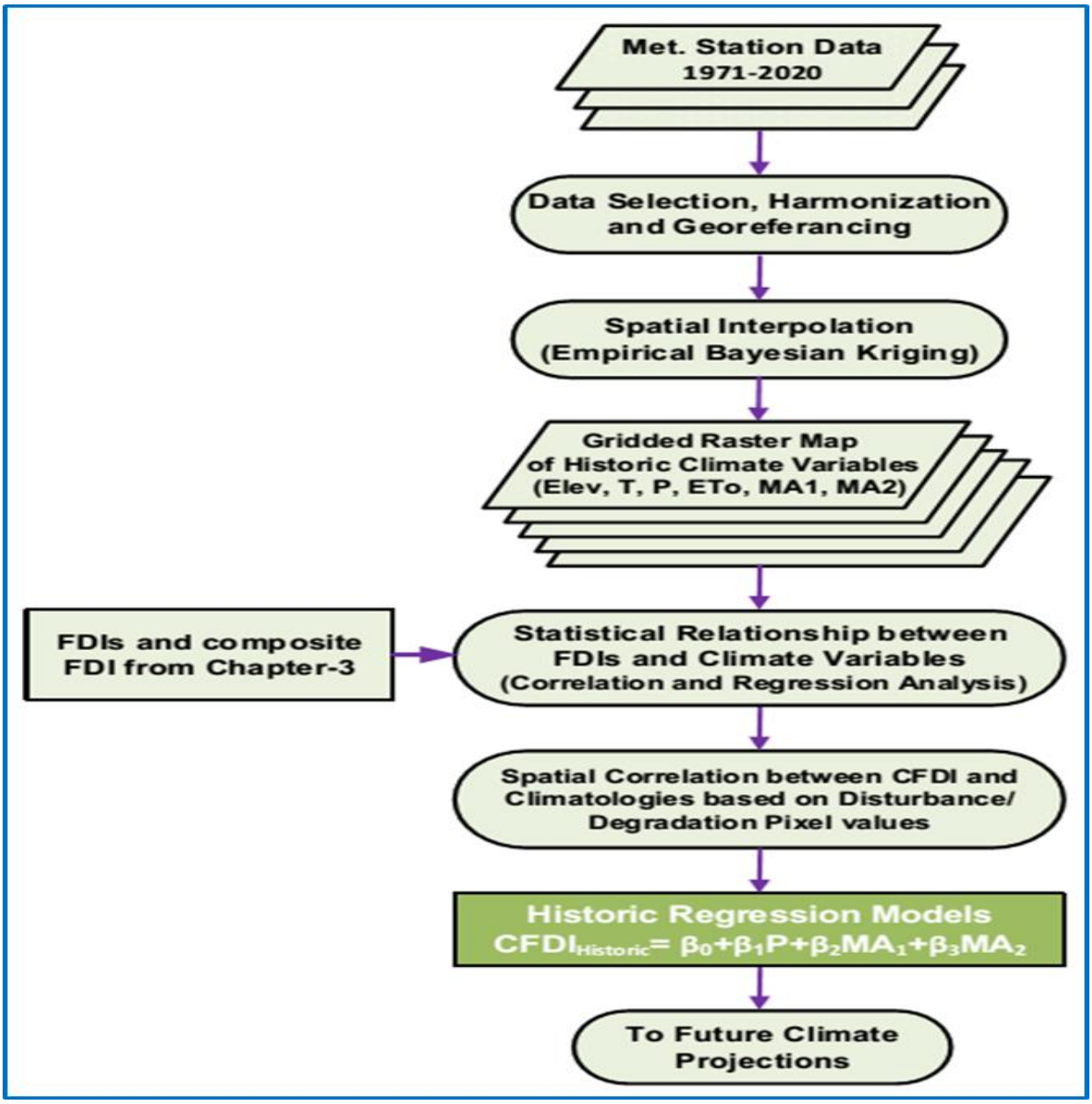
Zheng, Y., Han, J., Huang, Y., Fassnacht, S.R., Xie, S., Enze, Lv., and Chen, M. 2018. Vegetation response to climate conditions based on NDVI simulations using stepwise cluster analysis for the Three-River Headwaters region of China. Ecol. Indic., 92, pp.18–29.

Zhou, Z., Ding, Y., Shi, H., Cai, H., Fu, Q., Liu, S., and Li, T. 2020. Analysis and prediction of vegetation dynamic changes in China: Past, present and future. Ecological Indicators, 117, pp.106642.

Zontarelli, L., Dukes, M.D., Romero, C.C., Migliaccio, K.W., and Morgan, K.T. 2010. Step by step calculation of the Penman-Monteith evapotranspiration (FAO-56 method). University of Florida Institute of Food and Agricultural Sciences Doc. AE459.

Appendix 4

Methodological workflow of the relationships between climate variables and forest degradation



CHAPTER FIVE

PROJECTED FUTURE FOREST DEGRADATION INDICATORS UNDER CLIMATE CHANGE

Abstract

Global warming causing climate change, as a result of greenhouse gas (GHG) emissions, is predicted to significantly affect the frequency and severity of extreme weather events and ecological disturbances (e.g., forest fires). By the end of this century, on average, the forest area burned will likely double or even triple under the expected increases of the concentration of CO_2. In 2022, the Intergovernmental Panel on Climate Change (IPCC) warned North America is set to reach, or even surpass, the 1.5^0C level by 2050 if carbon emissions are not reduced. Ontario forests are already being affected by natural and anthropogenic disturbances (e.g., insect-induced damage, clear-cut harvesting) leading to forest degradation, which can significantly damage an ecosystem and impact a community's health, safety, and economy.

To ascertain how historic climate variables relate to forest degradation indicators (FDIs), the computed FDIs from available multispectral image stacks were statistically related to historic climatic variables, to used such relationships to project FDIs into the future with projected climate variables derived from General Circulation Models (GCMs). This research evaluated the usefulness of some FDIs in explaining patterns of forest disturbance and established quantitative relationships between useful FDIs and climate variables. Then, under the assumption such relationships with historical data will hold in the future, it employed these quantitative relationships to predict how future climate would determine the state of future FDIs in a designated protected area (Algonquin Park) in central and eastern Ontario, Canada. Signs of climate change impacts on Algonquin Provincial Park are evident. Trends in temperatures in and near the area are a reflection of the rise in daily average temperature, which was 4.9^0C, 5.1^0C, and 5.2^0C during the periods from 1971–2000, 1981–2010, and 2011–2020, respectively.

This study focused on an analysis of the forest dynamic changes in the Algonquin Park study area from 1972–2020 and generating predictions of possible changes in 2050. The effects of temperature (T), precipitation (P), evapotranspiration (ETo), and moisture availability (estimated from two regimes: MA_1 and MA_2) changes on forest vegetation conditions were examined from the perspective of the effects of temperature and water deficit. To determine the quantitative relationships between climate variables and FDIs, Pearson correlation coefficients and multiple linear regression (MLR) analysis were used for predicting future dynamic changes in forest cover. These quantitative relationships were used to compute forest degradation projection maps, which revealed forest cover will likely experience losses and significant fluctuations in the future, thereby putting the area's ecosystems at risk.

Moreover, in a parallel line of enquiry in this study, alterations in forest cover (forest degradation) were monitored from 1972–2020. These were derived from the composite forest degradation indicator (CFDI) computed from the multivariate (Principal Component Analysis [PCA]) resulting from the integration of seven FDIs. The relationships between remote sensing indicators (CFDI, NDVI, and NPP) and climate variables (T, P, ETo, MA_1, and MA_2) resulted in the creation of spatially-explicit maps of relevant climatic variables as well as from the long-term historical forest degradation maps (created by the trajectory analysis computed with LandTrendr) (Kennedy et al., 2010).

Forest cover changes under different climate change scenarios were subsequently evaluated based on both projected forest degradation scenarios resulting from the regression equations between indicators and the projected climate variables. Historical and future forest cover maps were produced to identify the areas with projected high vulnerability to climate change given their forest cover type. Finally, the difference maps of historical (1971–2010) and projected future (2050) forest cover were generated to discover potential changes in forest cover and vulnerable areas in the future, according to the projected climate.

An MLR model for the CFDI was fitted to historic climate variables and applied to the projected (2050) climate variables under the representative concentration pathway 8.5 (RCP8.5) scenario. Based on the differences between the historical (past and present) and future (year 2050) CFDI map, a future forest cover map was produced. The results predicted forest cover reductions, mostly in the south and southeastern portions of the study area.

The findings also indicated what climate variables (P, MA_1, and MA_2) had the strongest correlation (-0.95, -0.91, and -0.93, respectively) with the CFDI. This strong relationship allowed for computing an MLR model with a high coefficient of determination, R^2 (0.93), and low RMSE (0.28). This model was used for projecting the CFDI into the future (2050). A comparison of the historical average of 50 years (1971–2020) of climate normals (1971–2010), including average daily climate data (2011–2020), indicated that 2050 will experience daily average temperature and evapotranspiration increases (3.0°C, 0.2 mm, respectively). In addition, there will be a decrease in daily average precipitation (-0.3 mm) and the projected daily average moisture deficits, MA_1 and MA_2, will be 0.078 mm and 0.103 mm, respectively. It is also projected that the yearly decrease in precipitation will be 109.5 mm, the ETo increase will be 73.0 mm, and the moisture deficits will be 28.47 mm (MA_1) and 37.60 mm (MA_2) respectively. All these factors will result in increased forest degradation as determined by the FDIs.

Keywords—Climate change impacts, Ontario forests, forest disturbance and degradation, forest degradation indicators, LandTrendr.

5.1 Introduction

5.1.1 Climate change

The World Meteorological Organization (WMO) defines climate as the average weather over an extended time period (for instance, three decades). Climate change can be considered a shift over the long term (decades) in the weather on a global scale or for a particular place and time of the year resulting directly

or indirectly from anthropogenic activities and that can be identified by changes in the mean and/or variability of climatic properties (IPCC, 2013). A region's climate is the result of a complicated and interactive climate system with five main components: the atmosphere, hydrosphere, cryosphere, land surface, and biosphere. There are many biological, physical, and chemical processes involved in these interconnected elements. The climate is variable over space and time. In the span of the past 4.5 billion years, the Earth's climate has gradually changed.

In the past few decades, a great deal of progress has been made in understanding the climate system. It is clear now that anthropogenic factors, especially following the advent of industrialization, affecting atmospheric composition, are the primary actors responsible for the recent global warming trend. This is revealed by the fact that since industrialization in the 19^{th} century, climate change has occurred at an increasing rate, in particular after 1970. Recently, the rapid nature of climate change has become apparent through rises in global mean temperature and global mean sea level (GMSL), the alarming rate of ice melt around the globe, the increased frequency of extreme weather events, as well as changes in sea water acidification, precipitation, atmospheric moisture, and oceanic and atmospheric circulation patterns (IPCC, 2013).

Since 2001, 16 of the 17 warmest years on record have been experienced (NASA, 2018). In 2016, it was 0.99^0C warmer when compared to the average for the approximately 30-year period from 1951–1980. The Intergovernmental Panel on Climate Change (IPCC) recently (2022) warned that North America will likely reach, and even surpass, the 1.5^0C level by 2050. Without a reduction in GHG emissions (e.g., CO_2), the warming trend will continue, resulting in predicted rises in global temperature ranging from approximately 1.5 to 5.5^0C for the rest of this century (IPCC, 2014a). A synopsis of climate change projections for Ontario, Canada, indicated a 4.7 to 7.9^0C rise in average annual air temperature will occur in the basins of Lake Huron and the Ottawa River, when compared to the record of the years from 1971–

2000 (McDermid et al., 2015). Climate warming is expected to lead to longer growing seasons (1–30% greater in length by 2100), as well as increased forest fire frequency and pest outbreaks in Canada. By reducing and mitigating carbon emissions, warming can be slowed down and possibly the consequences to humanity and the environment could be avoided.

By the end of this century, in comparison with 1986–2005, the IPCC (2013) has predicted an increased global surface temperature from about 0.3^0C (in a low emissions scenario) to 4.8^0C (in a high emissions scenario). The appearance and functioning of ecosystems, by the end of this century, may also be very different when compared to today. Since the study area, which includes Algonquin Park (a designated protected area) is very important in conservation and public interest, it is crucial to have the information provided by updated climate warming projections at a scale useful for park management. These will help overcome the previous disparity between climate model projections at very coarse scales and potential effects at fine scales in terrestrial and aquatic ecosystems (Maclean et al., 2017; Nadeau et al., 2017).

5.1.2 GHG emission scenarios

The Earth's atmosphere is mainly composed of oxygen and nitrogen gases, with greenhouse gases (GHG) making up only a small part of the atmosphere. However, GHGs have a profound effect on the climate since they act as a radiation blanket that plays a role in warming the surface of the Earth. This "greenhouse effect" ensures our planet is warm enough for habitation by humans and other species. The accumulation of GHGs in the atmosphere (IPCC, 2007) causes global warming. This allows shorter wavelength incoming solar radiation to reach the surface of the planet while absorbing longer wavelengths (outgoing terrestrial), which is, in part, radiated back to the planet. In Earth's atmosphere, water vapor is the most abundant GHG, but its atmospheric concentration is moderated by temperature and not directly affected by anthropogenic activities. Unfortunately, a series of other GHGs—chlorofluorocarbons (CFCs), carbon dioxide (CO_2), nitrous oxide (N_2O), methane (CH_4), and ozone (O_3)—are strongly associated with human

activities. CO_2 is of special significance due to its long life and great abundance. GHGs in the atmosphere come primarily from burning fossil fuels (80%), while changes in land use and land cover, especially forest degradation and deforestation, contribute the rest (20%) (Stern, 2007).

Over the last few decades, a significant amount of evidence related to climate change had been compiled. With the goal of estimating future emissions and atmospheric concentrations of GHGs, IPCC Working Group I has created long-term emission scenarios, now based on representation concentration pathways (RCPs), instead of scenarios based on development pathways. These scenarios, however, omit natural emission sources, such as volcanic eruptions. It is the goal of such scenarios to provide a description of how radiative forcings have the potential to influence the future emission scenarios of energy balance between radiative energy absorbed by the planet and that radiated to the outer atmosphere. These are the RCPs that are dependent on GHG concentrations in the atmosphere. By employing the approximate total radiative forcing in 2100 compared to that in 1750, four RCPs (2.6 Wm^{-2} for RCP 2.6, 4.5 Wm^{-2} for RCP 4.5, 6.0 Wm^{-2} for RCP 6.0, and 8.5 Wm^{-2} for RCP 8.5) are detailed in the IPCC Fifth Assessment Report (AR5) (IPCC, 2013). These RCP categories differ depending on when carbon emission reductions are implemented (i.e., early, mid-century, late, or not at all). These categories are RCP2.6 (no longer used due to the timeframe of 2010–2020), RCP4.5 (CO_2 peaks in 2040 then declines), RCP6.0 (CO_2 peaks in 2060 then declines), and RCP8.5 (status quo continues with CO_2 continuing its upward trajectory at the current rate). This study used RCP8.5 for 2050 (GHG emissions continue at their current rate and CO_2 concentrations continue to increase) in order to have projections of outcomes if no mitigation at all occurs (worst-case scenario).

Considering the essential role vegetation plays in material, energy, and carbon balance; the hydrological cycle; the reduction of GHGs; and climate stability (Zhang et al., 2013), understanding vegetation responses to climate change is crucial to the survival of all life on Earth. Many recent studies in different

ecosystems (Jiapaer et al., 2015; Ren et al., 2020) have shown how vegetation dynamic changes are closely related to climatic variation. So, the analysis and prediction of vegetation dynamic changes becomes imperative, especially with the increasing influence of climate change (Fu et al., 2010).

Obtaining remote sensing data for studying the dynamic changes in vegetation is fortunately easier today than in the past due to progress in remote sensing technology (Huete, 2016). In addition, the fast pace of evolution in computer technology and popularization of remote sensing data have facilitated the large-scale monitoring of environmental phenomena (Shi and Chen, 2018; Wang et al., 2019). The most important indicator in remote sensing, the normalized difference vegetation index (NDVI), has been frequently employed as the basis of detection of vegetation photosynthetic activity, enabling the examination of variables such as vegetation coverage, vegetation condition index (VCI), vegetation productivity, leaf area index (LAI), net primary productivity (NPP), normalized burn ratio (NBR), enhanced vegetation index (EVI), and tasseled cap transformation (TCT). Climate change is the primary instigator of vegetation changes (Gu et al., 2018; Zhang et al., 2020). Additionally, ETo, which functions as another basic part of the hydrological cycle and is related to water and energy exchange between the Earth's surface and its atmosphere, has the capacity to determine the energy and water transport in the crucial soil-vegetation-atmosphere system (Shi et al., 2017). ETo impacts vegetation dynamic changes (Sun et al., 2012; Zhang et al., 2019) and is involved in the moisture balance that assists in ecosystem services and water resource management. The relationships between ETo and NDVI should be explored, including other vegetation indices/indicators, to ensure the dynamic monitoring of vegetation changes.

5.1.3 Assessment of potential impacts under a changing climate

5.1.3.1 Assessment process of climate change impacts

Temperature and precipitation are the most relevant meteorological variables when it comes to climate impact studies (Bronstert et al., 2007). However, modelling precipitation is more complicated than temperature, primarily due to its nonlinear nature, and its high level of spatial and temporal variability.

Precipitation patterns and the frequency of extreme events are both affected by higher temperatures and increasing evapotranspiration (Trenberth, 2011). Climate change effects on regional water resources are evaluated to create future climate scenarios acquired from general circulation model (GCM) simulations. These are highly complex multi-layer models of the physics of the atmosphere, over course scales covering the entire planet at finite time steps, against the input of a range of GHG emission scenarios.

The IPCC both collects and reviews GCMs as a component of the International Climate Change Assessment Report (IPCC, 2013). The main function of GCMs is to comprehend the dynamics of the physical aspects that make up the climate system (ocean, atmosphere, land and sea ice) and to create projections based on future GHG and aerosol forcings. When these models from different sources are combined through "experiments" of model inter-comparisons, they are called the CMIP (Meehl et al., 2000). A framework for coordinating climate change experimentation is provided by CMIP phase 5 (CMIP5) (Taylor et al., 2012). Over the last several years, it has provided novel information about the climate system and the process that leads to climate change and variability. There are approximately 20 modelling groups that contribute to the CMIP5 simulations through over 50 GCM models (Taylor et al., 2012). GCM outputs are calculated under four main RCPs representing potential future emissions and concentration scenarios (Moss et al., 2010). Each RCP (8.5, 6.0, 4.5, and 2.6) furnishes the amount of radiative forcing expected to be created by GHGs in 2100 in units of Watts per meter squared (W/m2). RCPs have the benefit of providing the climate modelling community with a suitable foundation for exploring a wide range of potential climate outcomes. RCP8.5 is characterized by rising GHG emissions over the course of time, which will likely result in the equivalent of more than 1370 ppm of CO_2 by 2100 (Van Vuuren et al., 2011).

GCM and Regional Circulation Models (RCM) projections are employed for the downscaling of GCMs (GCM–RCMs couples). According to existing 21st century GHG scenarios, climate change should result

in elevated air temperatures. However, what precipitation changes might occur in the future are highly uncertain for many regions and estimates vary in response to the GCMs employed (Ma and Xie, 2013). In order to formulate estimations about and determine the probability of potential future scenarios, impact studies employing the outputs of a wide variety of GCMs or GCM–RCM couples are necessary.

5.1.3.2 Potential impacts of drought under a changing climate

Determining the length of growing period (LGP) of vegetation is a commonly used procedure for assessing the impact of climate change. The LGP is the period (expressed in days) during a year when precipitation is greater than half the potential evapotranspiration (P>0.5ETo) and ambient temperatures are above the "dormancy period" (6.5 ^{0}C). A period of this type fully meets the ETo demands of vegetation and replenishes soil moisture. Plants' LGP can be significantly increased or decreased as a result of climate change, particularly concerning not only moisture balance but the expected increases in air temperature above the dormancy period. There are also potential positive and negative impacts of changes in growing season length.

In order to develop a detailed picture of the effects of climate shifts on the growth and development of vegetation, LGP should incorporate other elements of projected climate change (e.g., temperature changes). Remote sensing based estimates of meteorological variables have been increasingly used due to their availability and easy access. Kuo et al. (2005) formulated a new process by applying remote sensing data for the estimation of evapotranspiration employing the Penman-Monteith method.

For the assessment of the future of forested ecosystems in the study region, it is important to elucidate whether climate change will result in drier conditions, as well as how these conditions might affect growth and disturbance in forest ecosystems. In particular, forests experiencing drought-stress are likely to be more vulnerable to damage by diseases and insect infestations, especially those that thrive in warmer temperatures (Volney and Hirsch, 2005). Computer-generated GCMs of climate change projections

generally agree that significant warming will occur. However, the predictions for moisture regimes are particularly uncertain. Even with slight increases in precipitation, could still see drier conditions as a result of warmer temperatures that increase evapotranspiration.

Drought is the main actor constraining tree growth and distribution in a landscape with limited water; however, fire is a potential agent of rapid change at the landscape level. Throughout Canada, the fire regime has already shaped the composition of the boreal forest with a gradual decrease in fire-intolerant balsam firs from the wet eastern boreal extending all the way to the much drier western boreal forest. While it is clear that climate change presents serious concerns for the future of forest productivity in this area, it is possible some negative consequences would be offset by the positive scenarios of warmer soils, longer growing seasons, and the CO_2 fertilization effect.

5.1.4 Ontario forests and climate change

Forests play a key role in maintaining the delicate relationship between natural ecosystems. As the climate and forests are interrelated, they affect each other directly and indirectly. However, forests are expected to face significant pressure in the coming decades from rising GHG concentrations and climate change (IPCC, 2013), potentially disrupting the important ecological, economic, social, and aesthetic services they provide (Eastaugh, 2008).

In Canada, there has been a significant, although variable, warming trend since the end of the Little Ice Age (1850), which seems to be increasing (Environment Canada, 1992). Researchers established that the incidents of forest fires have risen dramatically over the past 40 years (Duinker, 2008). Flannigan et al. (2005b) projected that by the end of this century, on average, the area burned in Canada will increase by 74% to 118% in a $3\times CO_2$ scenario. Ontario, a large and diverse geographical area that experiences marked climatic variations, has seen a rise in surface air temperatures from 0.5–0.7°C since 1895 (Gullett and Skinner, 1992). Global warming of 1–3.5°C is predicted over the century, which will result in changes to

climate and weather patterns that will greatly affect forest ecosystems, in particular through an increase in fire and insect activity (Cohen et al., 2016).

The boreal regions, such as those in the study area, are expected to undergo an increase in temperature of 4–10°C over the next 50–100 years as the planet's atmosphere experiences a doubling or even tripling of atmospheric CO_2 concentrations (IPCC, 2007). The largest portion of Canadian boreal forests is in Ontario (Ontario Ministry of Natural Resources, 2008b), where, since the mid-20th century, forests have been impacted by human activities, which are projected to change the composition and function of ecosystems significantly (Schindler, 1998). At present, a shift is being registered in the boreal forest from a net carbon sink to a net carbon source, primarily as a result of the increased intensity, area, and frequency of fires resulting from increased temperatures (Flannigan et al., 2005a, 2005b; Soja et al., 2007). The expected impact of climate change in boreal forests could be increases in ignitions from lightning, fire season length, and fire weather severity (Soja et al., 2007).

The most important effects of climatic change in Ontario forests are likely to be changes in forest disturbance regimes (e.g., fire and insect outbreaks). Using climate projections from GCMs and derived RCMs, studies have extrapolated how climate change and variability may impact the trends of lightning, temperature, precipitation, fuel moisture, and other relevant meteorological parameters that could increase fire occurrence and forest degradation. The size of the area burned, fire season length, and severity of boreal fires will likely increase due to climate change (Flannigan et al., 1998; Flannigan et al., 2001, as cited in Amiro et al., 2003) and greatly threaten life and property. Potential burn in the Ontario boreal forest could double or even triple by the end of the century, producing global implications by releasing GHGs, which will alter the energy balance on the Earth's surface. As this social-ecological system is vitally important and susceptible to climate change, it is imperative to identify and map effective methods to assess forest disturbance and degradation, as well as strategies for mitigation and adaptation.

5.1.5 Climate change and forest disturbance and degradation

Forest ecosystems, in particular, are tightly bound with climate both directly through the effects of precipitation, as well as temperature, and indirectly through the effects of disturbances. Forest disturbances result in changes to the composition, structure, and character of a forest ecosystem (Jonson and Miyanishi, 2007), affecting its future health (Duinker, 2008). Natural disturbances, such as weather, geological forces, or biological changes, are one reason for an unbalanced ecosystem. These disturbances, which should increase in magnitude with a changing climate, can be caused by fire, insects, diseases, severe wind and ice storms, floods, volcanic activity, earthquakes, droughts, long-term freezing, hurricanes, landslides, etc., and can significantly damage an ecosystem, even killing flora and fauna. Over the past 20 years, studies have shown that disturbance regimes are changing rapidly (Turner, 2010). Projections for the future have revealed that fires are likely to become more frequent and severe, leading to a change in the relationships among fire, climate, and vegetation (Lutz et al., 2009; Westering et al., 2011). Recently, more research has been conducted on interactions among disturbances; however, understanding interactions is still a challenge from the perspective of ecology (Turner, 2010).

Ecosystems are also affected by human activities (Cohen et al., 2016), which are the cause of most forest disturbances, including clear cutting and the conversion of forest to other land uses, such as commercial and residential development, highways, and agriculture. Other human disturbances include uncontrolled wildfires, primarily caused by escaped or deliberately set fires, the introduction of non-native species, the elimination of species, and climate change. Ontario's forests are significantly influenced by these natural and human-made disturbances. In the Ontario boreal forest, in particular, insect outbreaks, forest wildfires, permafrost melting, and logging could potentially worsen with global warming. Additionally, industrial development could weaken both the forest's resistance and resilience to the intensifying negative consequences of global warming. Slash-burning ("prescribed fire"), a well-established practice in Ontario

(Lawson, 1990), exemplifies some of the harmful effects of anthropogenic activities. As well as producing smoke and GHGs, it may have other negative repercussions, such a possible reduction and destruction of the long-term nutrient supply, micro-organisms, and organic matter in the soil (Fuglem, 1990).

Climate change, including the rise in extreme weather events, such as the increased frequency, duration, and intensity of drought as well as summer heat waves, hurricanes, severe thunderstorms, windstorms, and ice storms (Duinker, 2008), will have a significant effect on the function, composition, and structure of forest ecosystems (Soja et al., 2007; Wulder et al., 2007). The rising intensity and prevalence of insect disturbances will also greatly impact forests (Wulder et al., 2007; Duinker, 2008). On a provincial scale, in Ontario, the supply of traditional forest products is expected to decline, accompanied by a negative impact on the forest industry and forest-dependent communities (Browne and Hunt, 2007). Climate change also threatens existing areas that protect specific natural features, species, and ecological communities (Scott and Lemieux, 2005). Besides the effect on biodiversity conservation, such changes can have dire consequences on nature-based recreation and tourism (Browne and Hunt, 2007).

5.1.6 The selection of climate change projections: The ensemble model approach

Globally, for climate change projections, combining models into an ensemble furnishes a method of 'average' projected change that might occur as a result of the rise in GHG emissions. By downscaling GCMs to the regional scale (RCM), climate change projections are more relevant to human activity, plant life, and animal species. This is useful for planning purposes and to better comprehend how climate warming impacts ecosystems.

5.1.7 Bias correction of downscaling

Interactions among various elements—ocean, cryosphere, atmosphere, and land surface—can be simulated in the new generation of GCMs. On a global scale, these are presently the optimal modelling tools for predicting the long-term climate system responses to gradual GHG increases. Owing to the

development of greater computing power recently, GCM grid resolution has been improved from a few hundred to about 100 km (10 km X 6 km). A variety of downscaling methods have been developed for transferring climate data from the coarse-scale GCM outputs to scales that are regional or local. For the downscaling of GCM information, statistical and dynamical downscaling are mainly employed; however, there are other basic methods (e.g., interpolation) that have been utilized (Fowler et al., 2007).

Unfortunately, historical climate bias results greatly affect the sensitivity of future climate change projections using GCM simulation (Liang et al., 2008). When a GCM reasonably simulates conditions for the current and historical climate, its credibility will be higher for future climate projection. This is possible if a downscaling model is able to provide value to historical, present, and future GCM projected climate variables.

There is no good representation for the process governed from the regional to local scale in GCMs (Volosciuk et al., 2015). Additionally, given the significant biases in GCM simulations, they are patently unsuitable for direct use when trying to determine climate change impact at the local level. At this scale, one must complete downscaling and/or bias corrections prior to conducting an analysis of the effects of climate change (Onyutha et al., 2016). Owing to the fact that they employ physical or statistical methods, downscaling and bias correction attempt to remove as much small-scale information from GCMs as possible (Bürger et al., 2012).

5.1.8 Projected future climate scenarios

A future climate scenario can be defined as the predicted rise of carbon in the atmosphere that results from factors that include population and economic changes, the deployment of alternative energy sources, and new technology distribution around the world. Given the many different factors that can affect the Earth's climate, a wide variety of future climate scenarios exist (IPCC, 2007). Analyzing these potentialities is made more complicated by many different climate models for estimating temperature and precipitation

changes based on the various scenarios. Each one will lead to different results although they might employ the same data related to future atmospheric carbon conditions. Therefore, a myriad of possible values for temperature and precipitation exists depending on the climate scenario and model selected. For instance, the assumption that the atmospheric CO_2 concentration will reach 850 ppm in 2100 (IPCC, 2007).

With the updating of model scenarios (IPCC, 2014b), our certainty related to the weight of evidence has improved (Cook et al., 2016), and the ability to downscale GCMs has ameliorated climate projections on the regional scale. Being able to project climate effect in the study area at a grid scale of 10 X 6 km allows for the evaluation of change in the region. Since the Algonquin Park landscape includes various forest types, growing conditions, and topographic scenarios, conducting a comprehensive assessment of change inside its boundaries (and the surrounding area of study) is a crucial step in comprehending patterns of climate change affecting all ecosystems in the study area.

Using present conditions, the multiple linear regression (MLR) model was fitted to data in order to project and identify the most vulnerable forest degradation areas under the climate change scenario. In order to achieve this objective, a composite forest degradation-based indicator (CFDI) was developed to determine the relationship between climate variables and forest cover under the current conditions. The data for 1971–2010 (average for a 40-year period) were processed and a multiple regression equation was created using forest degradation indicators (FDIs) as the dependent variables and climate variables as the independent variables.

Forest cover changes under the climate change scenario (RCP8.5) were subsequently evaluated based on both the regression equation and projected climate variables. Further, the historical, present, and future (under the RCP8.5 climate scenario) forest cover maps were produced to identify areas that are vulnerable in terms of their degradation of forest cover. Finally, the difference maps of historical (1971–2010) and

projected future (2050) forest cover were generated to map out potential changes in forest cover and vulnerable areas in the future.

In this chapter, vegetation dynamic changes were analyzed for 1972–2020, and possible changes in 2050 were examined for the study area. The effects of temperature, precipitation, and ETo on moisture availability (MA_1 and MA_2) in forests from the perspective of water deficit were considered, with MLR analysis being used as the foundation for creating projection models. By employing these projection models and the future meteorological data from CMIP5 models, the future dynamic changes of forest vegetation were predicted for the Ontario study area. In particular, the goal of this work was to (i) predict the expected change in forest condition (degradation) with climate change in the forest area for the region in the future (e.g., year 2050), and (ii) determine the consequences of projected forest degradation with climate change in the region of study relative to GHG concentrations and fluxes. This will hopefully serve as the basis for the development of appropriate climate change mitigation and adaptation strategies for forest ecosystem and biodiversity management for the region in the future.

5.2 Materials and methods

5.2.1 Methodological workflow

In this research, projected meteorological data provided by the IPCC were employed. The IPCC published a series of four emission scenarios (i.e., RCP2.6, 4.5, 6.0, 8.5) in 2013, each representing a different GHG emission concentration influence by anthropogenic activities (IPCC, 2013; Shi et al., 2018). This study chose only the future meteorological data under the RCP8.5 high-emissions “worse-case” scenario. The multi-models/ensemble average was used to evaluate the future changes in an effort to achieve more reliable results. For the purposes of unifying the spatial resolution of the VIs, the cell sizes of the meteorological variables were resampled to 30m resolution using the robust EBK interpolation method, which is superior to all other interpolation methods.

Although it is common for models to be comprised of multiple ensemble members, for the Ontario study area, only the first member from each model ensemble was employed. As a result, each model used in this study was given equal weight for the performance of simple model statistical analysis on the downscaled ensemble (Moss et al., 2010; Flato et al., 2013). The period 1971–2010 was taken as the reference period for the bias correction. Changes in the year 2050, relative to historical conditions, were investigated. Accordingly, results garnered from precipitation, temperature, and some derived variables were analyzed and presented for these periods (Table 5.2).

The future climate projection workflow methodology was constructed in three main steps (Figure 5.1): (i) Forest degradation map (long-term historic trajectory) (details in Chapter 3 and 4); (ii) Gridded raster map (historic climate variables) (details in Chapter 4); and (iii) GCM data (downscaled to RCM data) using an ensemble of 24 GCMs.

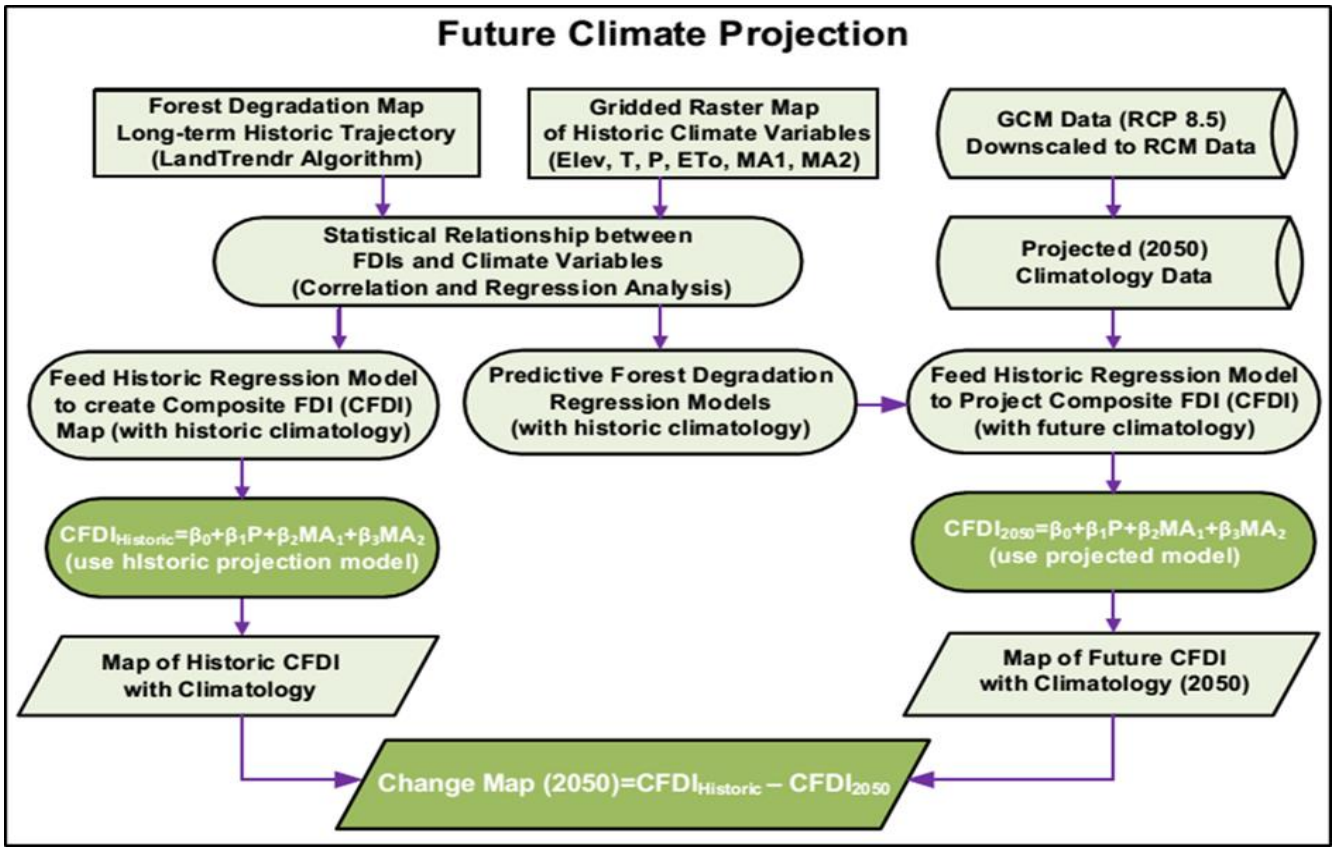


Figure 5. 1: Methodological workflow for projected future forest degradation under climate change

5.2.1.1 Forest degradation map from long-term historic trajectory

In this study, seven FDIs—normalized difference vegetation index (NDVI), net primary productivity (NPP), normalized burn ratio (NBR), enhanced vegetation index (EVI), tasseled cap brightness (TCB), tasseled cap greenness (TCG), and tasseled cap wetness (TCW)—were calculated from the satellite images (Landsat and MODIS platforms) using ENVI^{TM} software. All FDI variables were combined into a CFDI, using the image fusion technique reliant on principal component analysis (PCA) (details in Chapter 3).

A long-term (1972–2020) historical forest degradation trajectory/trend map was created by the "trajectory analysis procedure" computed from the LandTrendr algorithm (Kennedy et al., 2010). A correlation matrix was developed based on the pixel values of the degradation area for 1972–2020 as revealed by LandTrendr results to examine the relationship between indicators derived from remote sensing and

climate variables. For the future projections, regression analysis was conducted based on the highly correlated dependent (CFDI) and independent (P, MA_1, and MA_2) variables (details in Chapter 4). The structure of the MLR model is as follows:

$$CFDI\ (Y) = \beta_0 + \beta_1 P + \beta_2 MA_1 + \beta_3 MA_2 + \varepsilon \quad (5.1)$$

Another correlation matrix was created based on all climate variables—including CFDI, NDVI, and NPP—for 2001–2020 to examine the correlations between indicators derived from remote sensing and climate variables. For the future projection, regression analysis was conducted based on the highly correlated dependent variables (CFDI, NDVI, NPP) and independent variables (P, MA_1, and MA_2) (details in Chapter 4). The structures of the MLR models for NDVI and NPP are as follows:

$$NDVI\ (Y) = \beta_0 + \beta_1 P + \beta_2 MA_1 + \beta_3 MA_2 + \varepsilon \quad (5.2)$$

$$NPP\ (Y) = \beta_0 + \beta_1 P + \beta_2 MA_1 + \beta_3 MA_2 + \varepsilon \quad (5.3)$$

Finally, these equations were implemented on P, MA_1, and MA_2 maps to spatially interpolate the values of meteorological variables over the entire targeted area at 30m spatial resolution using the EBK interpolation method.

5.2.1.2 Gridded raster maps of historic climate variables

Historical climate (daily and monthly) data for 1972–2020 were acquired from the official climate data sources of Environment and Natural Resources Canada. In addition, historical Canadian climate normals data were garnered from the Environment and Climate Change Canada (ECCC) official climate data portal. Two sets of 30-year climate normals data (1971–2000 and 1981–2010) were collected from 51 weather stations inside and surrounding the study area. The geostatistical modelling has as its foundation the spatial auto-correlation as depicted by the semi-variogram, which requires point data for as many locations as possible.

These data were compiled for ETo calculations for all 51 weather stations. ETo was calculated using the Penman-Monteith method (ETo Calculator version 3.2), furnished by the Food and Agriculture Organization of the United Nations (FAO-56) (Allen et al., 1998; Yao et al., 2018). ETo calculation was required for MA_1 (P-ETo) and MA_2 (P-0.5ETo) calculation (see Chapter 4 for details). While MA1 is a more demanding estimate of water balance due to its use of full evapotranspiration returns of moisture to the atmosphere, MA2 using half of the evapotranspiration has been calibrated and is more commonly used in the estimation of the length of the growing period (LGP) for vegetation growth, globally (FAO, 1978)

As previously stated in Chapter 4, EBK automates the most complicated feature in the construction of a valid semi-variogram model, which is also the aspect of spatial interpolation that is the most challenging within the geostatistical method (Gupta et al., 2017). The K-Bessel semi-variogram model was also selected since it is very flexible and accurate in furnishing spatial estimation of meteorological variables (e.g., T, P, ETo, MA_1, and MA_2).

The compiled data table (Table 4.3 in Chapter 4) was open in ArcGIS after completion of the calculations of ETo, MA_1, and MA_2 with historical climate normals data (average of 1971–2010) for 51 meteorological stations, and a point feature map was created. This feature map was applied to create the prediction/gridded raster maps of Elev, T, P, ETo, MA_1, and MA_2, using the EBK interpolation method in ArcGIS^{TM} software (see Figures 4.4, 4.5, and 4.6 in Chapter 4). A correlation matrix was developed based on the pixel values of the long-term (1972–2020) forest degradation trajectory map, and the relationships between FDIs derived from remote sensing and climate variables were examined. For the future projections, regression analyses were conducted based on the highly correlated dependent (CFDI) and independent (P, MA_1 and MA_2) variables (details in Chapter 4).

The projection model of the CFDI was constructed based on the MLR equation. For each raster map, the NDVI, NPP, and CFDI values were expressed as a function of P, MA_1, and MA_2. Therefore, using the

values of the indicators from 1972–2020 as the dependent variable and the climate variables during the period from 1971–2010 as the independent variables, the structure of the projected regression model is expressed as follows:

$$\text{Projected CFDI } (Y) = \beta_0 + \beta_1 P + \beta_2 MA_1 + \beta_3 MA_2 + \varepsilon \quad (5.4)$$

Where the dependent variable CFDI is the composite forest degradation indicator; the independent variables P, MA_1, and MA_2 are the climate variables; and β_0, β_1, β_2, and β_3 are the corresponding regression coefficients (ε is a residual).

5.2.1.3 GCM data (downscaled to RCM data)

Monthly observations of maximum temperature (T_{max}), minimum temperature (T_{min}), mean temperature (T_{mean}), and monthly total precipitation (P) for the year 2050 were acquired from the Canadian Centre for Climate Services (CCCS) data portal (https://climatedata.ca). These data were already calculated by the CMIP5 climate model dataset and (statistically) downscaled and bias adjusted using the "BCCAQv2" method under scenario RCP8.5. Bias-correction constructed analogues with quantile mapping BCCAQ is a hybrid downscaling method of bias-correction constructed analogues (BCCA-Maurer et al., 2010) and quantile mapping (QMAP; Gudmunsson et al., 2012). An ensemble of 24 GCMs was used.

As in the section above, these data were compiled and prepared using a spreadsheet for ETo calculations for all 32 weather stations inside and near the research zone. ETo was calculated using the Penman-Monteith method (ETo Calculator version 3.2), which was required for MA_1 (P-ET_0) and MA_2 (P-0.5ETo) calculation (details in Chapter 4). EBK, which automates the complex process of creating a valid kriging model, and the K-Bessel semivariogram model, which provides more accurate spatial estimation of the meteorological variables, were also employed.

The compiled dataset was open in ArcGIS after completion of the calculation of ETo, MA_1, and MA_2 with projected (2050) data from 32 meteorological stations, and a point feature map was created. Next, the

interpolated feature class (gridded map) was applied to create the future projection maps of the climate variables P, MA_1, and MA_2 for 2050. The future projected map of CFDI was produced using the future (2050) projection maps of the climate variables P, MA_1, and MA_2. The raster calculator tool was applied to calculate the projected map using the following projection MLR model:

$$CFDI\ (Y)_{2050} = \beta_0 + \beta_1 P_{(2050)} + \beta_2 MA1_{(2050)} + \beta_3 MA2_{(2050)} \quad (5.5)$$

5.2.2 Downscaling techniques

In principle, downscaling techniques can be used to refine any type of data (Rummukainen, 2010). Having a GCM output that is coarse might be satisfactory (e.g., when the variation in a single grid cell is low or for global assessment). While GCMs have value when used as predictive tools, they lack the ability to account for heterogeneity on a fine scale and are unable to reflect on certain types of features (such as bodies of water, mountains, infrastructure, land cover characteristics, breezes along coastlines, and convective clouds). Unfortunately, it is very challenging to address this disparity between climate model resolution and regional and local scale processes.

GCMs, which are efficient and trustworthy tools for global scale climate analyses, are developed on the basis of numerical representations of the climate system. They include physical, chemical, and biological processes/properties of climate variables and feedback relationships that exist between these variables. GCMs are able to create simulations of the present climate and offer predictions of future climate change with forcing by GHGs and aerosols. However, as GCMs provide global scale data, tools are needed for regional studies to translate this information to the local scale. Downscaling techniques are widely used for transferring coarse-scale climate information to regional and local scales. There are two different downscaling techniques: dynamical and statistical.

5.2.2.1 Dynamical downscaling

Dynamical downscaling (DD) is based on complex algorithms of the atmospheric radiative, energy, and thermodynamic processes. It is one method of downscaling large-scale GCM climate output into a finer spatial resolution by taking the outputs of the GCM and running the physics and other complementary software packages within a given GCM "boxel." DD relies on the nesting of a finer scale RCM, one that is as fine as 10km X 10km in horizontal resolution in GCMs (Wood et al., 2004). This type of downscaling has only been attempted using three approaches: (i) simulating a regional scale model using GCM information for geographical boundary conditions; (ii) running a global scale experiment that has high-resolution atmospheric GCMs, as well as coarse GCM data acting as initial and partial boundary conditions; and (iii) using a global model with variable resolution.

The goal of DD is the spatial and temporal refinement of climate data for a specified region engaging in the description of phenomena and forcings that are unresolved in GCMs. In order to achieve this objective, DD employs large-scale lateral boundary conditions with the aim of ensuring the GCMs are able to create higher resolution outputs through the nesting of an RCM in the GCM. In this process, RCMs parameterize and resolve atmospheric processes. As a result, they are able to provide realistic simulations of regional climate features for the production of higher resolution outputs (approximately 0.1 to 0.50 in latitude and longitude). These include orographic precipitation (e.g., Frei et al., 2003), climate events that are extreme in nature (e.g., Fowler et al., 2005; Frei et al., 2006), and climate anomalies at the regional scale. GCM–RCM couples (RCM outputs) even have the capacity to provide representations of non-linear effects, for instance those connected with the El Niño Southern Oscillation (e.g., Leung et al., 2003).

When it comes to a DD framework, however, it should be noted that RCM skills depend greatly on biases coming from the driving GCM and also the existence and strength of regional scale forcing (e.g., orography, contrast between land and sea, and vegetation cover) (Fowler et al., 2007). This means that

the substantial errors inherited from the driving GCMs are retained in the GCM–RCM couples (Rummukainen, 2010; Hall, 2014). For simulations related to temperature, the uncertainty from the RCMs may, in fact, be less than that arising from the emission scenario (RCPs); however, the opposite is true for precipitation simulations (Fowler et al., 2007). Their great dependency on GCM boundary forcing is one of the primary downsides of RCMs. Therefore, GCM–RCM couples generally are unable to accurately reproduce the average seasonality of precipitation or average annual precipitation for the historical period, something that has been revealed for various regions (Tramblay et al., 2013; Dakhlaoui et al., 2019; González-Zeas et al., 2019). This means that RCM raw outputs usually need additional statistical bias correction before they can be employed for hydrological model inputs. In short, the primary downsides of this downscaling method include model complications, high computational requirements, dependence on boundary conditions gleaned from GCMs, and a lack of transferability to other regions.

5.2.2.2 Statistical downscaling

Statistical downscaling (SD), which is also called empirical/statistical or statistical/dynamical downscaling (Mearns et al., 2003), has as its basis the concept that regional climate is primarily affected by two factors: the large-scale climate and local/regional features (topography, land use, land-sea distribution) (Fowler et al., 2007; Wilby et al., 2004). In contrast to DD, SD transforms large-scale GCM (or GCM–RCM couples) outputs into a much finer resolution. It employs an empirical relationship between predictors (in this case, large-scale GCMs simulated climate variables) and predictands (in this case, regional scale variables, such as precipitation). There are currently three SD approaches: (i) classification/weather typing methods (Hughes and Guttorp, 1994; Mehrotra and Sharma, 2005), (ii) regression/transfer function methods (Wilby et al., 2002; Chen et al., 2014), and (iii) weather generator (WG) (Wilks and Wilby, 1999; Srivastav and Simonovic, 2014; King et al., 2015). In addition, when SD is employed, three assumptions are made (Hewitson and Crane, 1996): (i) the predictor variables are

modelled in a realistic manner by GCMs, (ii) the empirical relationship presented is valid for any climatic conditions necessary (both stationary and non-stationary), and (iii) the predictor is a good representation of the climate change signal. Due to its low level of computational requirements, simple modelling structure, and the ability to make easy modifications for use at various locations, SD is more popular, adaptable, and flexible. However, while there is computational efficiency in the statistical approach, in addition to its ability to downscale the GCM output to the station-scale level using enough climate information, SD must be performed with a long series of reliable historical data for proper calibration.

The SD process involves two steps. First, statistical relationships between the GCM projections and local variables must be established with observed historical climate information. Next, employing the statistical relationships, the GCM projections are transformed to regional or local scales. Using SD, it is possible to simulate multiple outputs—maximum and minimum temperature, precipitation, relative humidity, solar radiation, and wind speed (Parlange and Katz, 2000)—at the same time. This is significant, especially when it comes to impact studies (Wilby et al., 2004).

The performance of this downscaling technique is dependent on the selection of the regional domain (Wilby and Wigley, 2000), something that is generally not taken into account in practice (Benestad, 2001), and on the regionalization methods (Wilby et al., 2004). In their study, Gutierrez et al. (2013) examined the performance of SD approaches generally employed for downscaling temperature (including MLR, analogue methods, weather typing techniques, and regression conditioned on weather types) in terms of their robust applicability in climate change research. They established a novel validation framework that exploits anomalous warm historical periods. Based on this framework, the authors concluded regression methods are the most appropriate for climate change studies. These methods outperformed the others as weather typing underestimated the temperature in moderately warmer conditions and analogue methods,

although they were the best for the reproduction of the observed distributions, they were responsible for significantly underestimating warm period temperatures when compared to the values observed.

For regression models, statistical linear or nonlinear relationships between observed local climate variables and the GCM outputs should be established (Wilby et al., 2002). Despite the fact that these methods are quite simple, their downside is the likely lack of a stable relationship between GCM outputs and observations (Wilby and Wigley, 1997). For weather typing, it is necessary to group local meteorological variables in terms of their relationship to different atmospheric circulation classes (Bárdossy and Plate, 1992; Von Storch et al., 1993). This allows for the close linking of local variables to large-scale global circulation. It should be mentioned, however, that its reliability is dependent on there being a stationary relationship between local climate and large-scale circulation, in particular when it comes to precipitation. The WG method has as its basis the perturbations of climate parameters according to the changes projected between current and potential climate change by GCMs (Qian et al., 2010; Wilks, 2010). As stochastic models, WGs can be employed to create synthetic weather series that have unlimited lengths and that can be consistent with the present and also possible future climate change (Wilks, 2010).

5.3 Results and discussion

Monthly average T_{min}, T_{max}, and T_{mean} and monthly total precipitation (P) were acquired from 32 weather stations in and around the research zone. The data were compiled and prepared to estimate the ETo (with the Penman-Monteith method), MA_1, and MA_2 for all 32 weather stations. Table 5.1 shows a summary of calculated results for the Algonquin Park East Gate station.

Table 5. 1: Summary of calculated results of ETo, MA_1, and MA_2 for the Algonquin Park East Gate station

Months	Lat	Long	Elev	T_{max}	T_{min}	T_{mean}	P	Daily P	ETo	MA_1	MA_2
January	**45.54**	**-78.29**	**397**	**-2.4**	**-12.2**	**-7.1**	**86.1**	**2.8**	**0.7**	**2.079**	**2.429**
February				**-1.5**	**-13.0**	**-7.3**	**74.4**	**2.7**	**1.0**	**1.655**	**2.155**
March				**3.8**	**-7.7**	**-2.0**	**68.0**	**2.2**	**1.5**	**0.693**	**1.443**
April				**11.2**	**-0.9**	**5.3**	**72.4**	**2.4**	**2.6**	**-0.186**	**1.114**
May				**19.9**	**6.3**	**13.2**	**89.4**	**2.9**	**4.3**	**-1.417**	**0.733**
June				**25.1**	**11.7**	**18.6**	**78.3**	**2.6**	**5.3**	**-2.691**	**-0.041**
July				**28.3**	**15.4**	**21.6**	**63.0**	**2.0**	**5.6**	**-3.568**	**-0.768**
August				**26.6**	**13.6**	**20.1**	**75.0**	**2.4**	**4.9**	**-2.482**	**-0.032**
September				**22.1**	**10.0**	**16.2**	**105.9**	**3.5**	**3.5**	**0.028**	**1.778**
October				**14.7**	**3.7**	**9.1**	**83.2**	**2.7**	**2.1**	**0.584**	**1.634**
November				**6.5**	**-1.3**	**2.5**	**99.8**	**3.3**	**1.0**	**2.326**	**2.826**
December				**0.2**	**-8.3**	**-3.8**	**89.4**	**2.9**	**0.7**	**2.185**	**2.535**
Average						**7.2**	**82.1**	**2.7**	**2.8**	**-0.066**	**1.317**

Projections of daily average ETo, MA_1, and MA_2 were also estimated for all 32 meteorological stations for 2050. Averages resulting from the summary tables of all individual 32 stations were compiled into the following table:

Table 5. 2: Calculated results of all projected climate variables of 2050 (GCM/RCM data) for 32 stations

Station's name	Lat	Long	Elev	T	P	ETo	MA_1	MA_2
AP Achray Campground	45.86	-77.75	228.0	7.7	2.4	2.841	-0.428	0.991
AP Basin Depot	45.71	-77.79	260.0	7.9	2.4	2.883	-0.503	0.938
AP East Gate	45.54	-78.29	397.0	7.2	2.7	2.766	-0.066	1.317
AP Sand Lake Gate	45.87	-77.54	184.0	8.1	2.4	2.85	-0.499	0.925
Arnprior	45.46	-76.38	107.0	9.1	2.4	2.841	-0.475	0.945
Bancroft	45.04	-77.88	348.0	8.1	2.6	2.875	-0.236	1.201
Barry's Bay	45.53	-77.67	289.6	7.9	2.4	2.875	-0.447	0.99
Combermere	45.29	-77.63	287.0	8.0	2.5	2.858	-0.394	1.034
Deep River	46.09	-77.49	142.0	7.9	2.5	2.858	-0.403	1.026
Eganville	45.52	-77.16	223.0	8.5	2.3	2.891	-0.581	0.863
Haliburton	45.04	-78.54	330.0	8.1	2.9	2.875	0.061	1.499
Huntsville	45.29	-79.21	320.0	8.3	3.1	2.758	0.317	1.696
Kemptville	45.04	-75.63	99.0	9.6	2.6	2.833	-0.224	1.192
Lake Traverse	45.96	-78.04	236.0	7.8	2.4	2.866	-0.426	1.006
Lavant Station	45.03	-76.70	274.3	8.7	2.5	2.833	-0.309	1.106
Madawaska	45.54	-77.96	316.0	7.4	2.5	2.866	-0.321	1.111
Maniwaki QC	46.37	-75.96	189.0	7.5	2.7	2.816	-0.085	1.323
Matawatchan	45.13	-77.14	393.0	8.3	2.5	2.85	-0.383	1.041
Mattawa	46.31	-78.70	186.0	7.7	2.6	2.725	-0.124	1.237
Mont-Laurier QC	46.54	-75.54	244.0	6.9	3.0	2.741	0.219	1.59
Muskoka	45.04	-79.63	282.0	9.0	3.0	2.741	0.25	1.621
North Bay	46.29	-79.46	370.0	7.9	2.8	2.616	0.149	1.457
Notre Dame QC	46.12	-75.63	213.0	7.4	2.9	2.75	0.157	1.532
Ottawa	45.46	-75.71	114.0	9.3	2.7	2.758	-0.069	1.31
Pembroke	45.86	-77.25	161.0	8.6	2.4	2.9	-0.541	0.908
Petawawa	45.87	-77.29	183.0	8.4	2.4	2.875	-0.523	0.913
Powassan	46.12	-79.29	274.0	7.7	2.8	2.646	0.109	1.432
RD Joachims QC	46.21	-77.71	137.0	7.8	2.5	2.816	-0.295	1.113
Renfrew	45.46	-76.71	130.0	8.8	2.3	2.883	-0.544	0.896
Shawville QC	45.62	-76.54	168.0	8.6	2.5	2.866	-0.393	1.039
Spotswood	45.87	-76.88	124.0	8.6	2.5	2.883	-0.403	1.037
Stonecliffe	46.21	-77.89	196.0	7.8	2.5	2.816	-0.327	1.081
Average				8.1	2.6	2.82	-0.242	1.168

Most of the weather station results (26 out of 32 stations for the study region in Table 5.2) showed negative moisture availability (MA_1), meaning there was a moisture deficit during the months from April to August along with the Algonquin Park East Gate station (Table 5.1, highlighted). The average results of these stations also revealed an overall moisture deficit (-0.242mm or 0.242mm/day and 88.33mm/year), even though for some stations, an increase in precipitation was also accompanied by a moisture deficit due to the rise in temperature resulting in higher evapotranspiration (ETo). It was found that the moisture deficits

projected for 2050 would likely be 24.1 mm in the Algonquin Park East gate area and 88.3 mm in the overall study region.

5.3.1 Trends of changing climate (1971–2020 and 2050)

Temperature trends (Table 5.3) within and near the area of study are a reflection of a significant rise in daily average temperature of 4.9^0C, 5.1^0C, 5.2^0C, and 8.1^0C during the periods from 1971–2000, 1981–2010, 2011–2020, and 2050 (projected), respectively.

Table 5. 3: Trend of rising temperatures from 1971–2020 and 2050 (Calculated summary results of climate variables)

Results summary	**T (°C)**	**P (mm)**	**ETo (mm)**	**MA_1 (mm)**	**MA_2 (mm)**
1971–2000 = 30-yr average	4.9	2.8	2.5	0.155	1.399
1981–2010 = 30-yr average	5.1	2.9	2.7	0.174	1.143
2011–2020 = 10-yr average	5.2	NA	NA	NA	NA
1971–2020 = 50-yr average	5.1	2.85	2.6	0.164	1.271
Average (2050)	8.1	2.6	2.8	–0.242	1.168
Difference	**3.0**	**–0.3**	**0.2**	**–0.078**	**–0.103**

Note: NA (not available): ETo, MA_1, and MA_2 were not calculated due to missing precipitation data for most of the stations from 2011–2020.

A comparison of the historical average of 50 years (1971–2020) of climate normals (1971–2010), including average daily climate data (2011–2020), indicated that 2050 will experience daily average temperature and evapotranspiration increases (3.0°C, 0.2 mm, respectively) on average. In addition, there will be a decrease in daily average precipitation (-0.3 mm), and the projected daily average moisture deficits MA_1 and MA_2 will be 0.078 mm and 0.103 mm, respectively. It is also projected that the yearly decrease in precipitation will be 109.5 mm, the ETo increase will be 73.0 mm, and the moisture deficits will be 28.47 mm and 37.60 mm respectively. All of this will result in increased forest degradation as determined by the FDIs.

Historical climate normals for 1971–2000 and 1981–2010 indicated an increasing trend of precipitation, and 2050 projected results showed a decrease in precipitation and increase in evapotranspiration (even

though an increase in precipitation was also accompanied by a moisture deficit due to the rise in temperature resulting in higher evapotranspiration).

5.3.2 Prediction raster maps of climatologies (2050)

The compiled data table (Table 5.2) was opened in ArcGIS after completing the calculations of ETo, MA_1, and MA_2 with projected (2050) data from 32 meteorological stations to create a point feature map. Based on the climate variables highly correlated with FDIs, the raster maps of predicted P, MA_1, and MA_2 were created using the EBK interpolation method in ArcGIS (Figure 5.2 a-c).

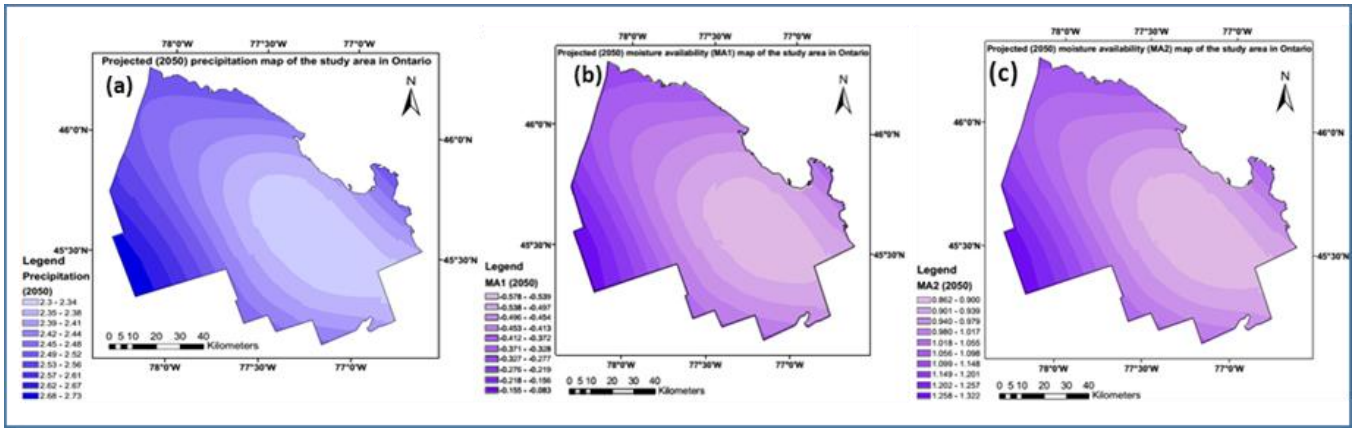


Figure 5. 2 (a-c): Prediction (2050) gridded maps of (a) P, (b) MA_1, (c) MA_2

The graduated single color ramp combination for the year 2050 showed daily average precipitation from 2.30 mm–2.73 mm. The prediction map (Figure 5.2a) indicated that the projected (2050) precipitation would likely decrease in most of the study area, except the southwestern section.

The graduated color ramp of light blue to dark blue for daily average MA_1 (–0.578 mm to –0.063 mm) and MA_2 (0.862 mm–1.322 mm) reveals the lower to higher moisture deficit of the prediction map (Figure 5.2b-c) and the daily average indicating a greater moisture deficit in the south, eastern, and middle portions of the study area.

5.3.3 Projected maps of climatology (2050)

An MLR model was developed based on the highly correlated dependent variable (CFDI) and the independent variables (P, MA_1, and MA_2) (details in Chapter 4). The interpolated feature class (gridded) map of 2050 was applied to create the future projected map. The MLR model for CFDI was fitted to the climate variables and applied to the projected (2050) climate variables under the representative concentration pathway 8.5 Wats/m^2 (RCP8.5) scenario. The 2050 projected model (based on 1972–2020 historical data) is as follows:

$$\text{Projected CFDI}_{\mathbf{2050}}\ (\hat{Y}) = 42.532 - 8.888\ P_{\mathbf{(2050)}} + 15.625\ MA_{1\mathbf{(2050)}} - 16.703\ MA_{2\mathbf{(2050)}} \quad (5.6)$$

The future projected map of CFDI was produced using the future (2050) prediction maps of the climate variables (P, MA_1, and MA_2). The raster calculator tool was employed to calculate the projected map (Figure 5.3a), fed by 2050 prediction climatic maps.

Another MLR model was developed (see Chapter 4) based on the highly correlated dependent variable, CFDI, and independent variables, P, MA_1, and MA_2, for 2001–2020, and the interpolated (raster) feature class maps of 2050 were employed to create the future projected map using the following projected model:

$$\text{Projected CFDI}\ (\hat{Y})_{\mathbf{(2050)}} = 17.0842 - 3.0950\ P_{\mathbf{(2050)}} + 2.0014\ MA_{1\mathbf{(2050)}} - 8.3937\ MA_{2\mathbf{(2050)}} \quad (5.7)$$

The CFDI projected model (5.7) was fed by the 2050 prediction maps of climatology, and the projected map was produced using the raster calculator tool in ArcGIS (see Figure 5.3b).

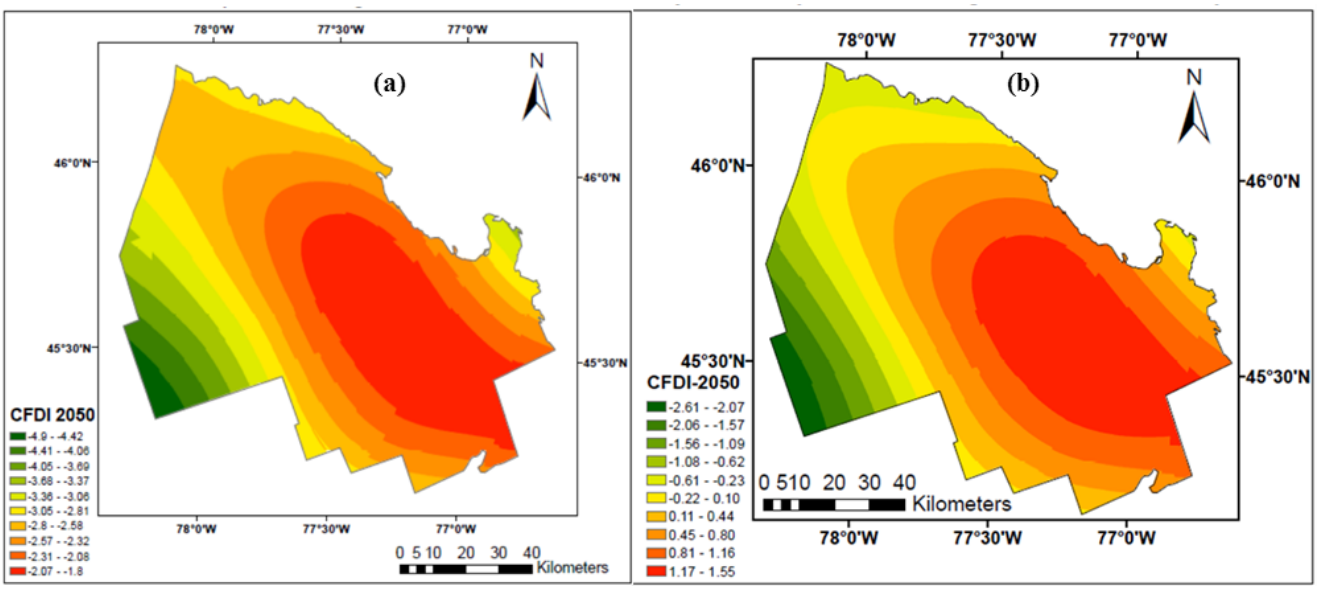


Figure 5. 3 (a-b): a) Projected (2050) CFDI map based on 1972–2020 model; (b) Projected CFDI map of 2050, based on 2001–2020 model

The colour-coded resulting maps indicate a high level of degradation in red; green denotes very low or no degradation. A lower level of degradation is represented by shades in between the green and dark red colors (Figure 5.3a–b).

5.3.4 Estimation of projected change map to 2050 of FDIs and mapping with climate change, relative to historic

In Chapter 4, two historical CFDI gridded maps were created using the predicted MLR models (Eq 4.5, 1972–2020 and Eq 4.6, 2001–2020). The models were fed by the historic prediction maps of climatology, and the calculations for the maps (Figure 5.4a-b) were once again undertaken with the raster calculator tool in ArcGIS.

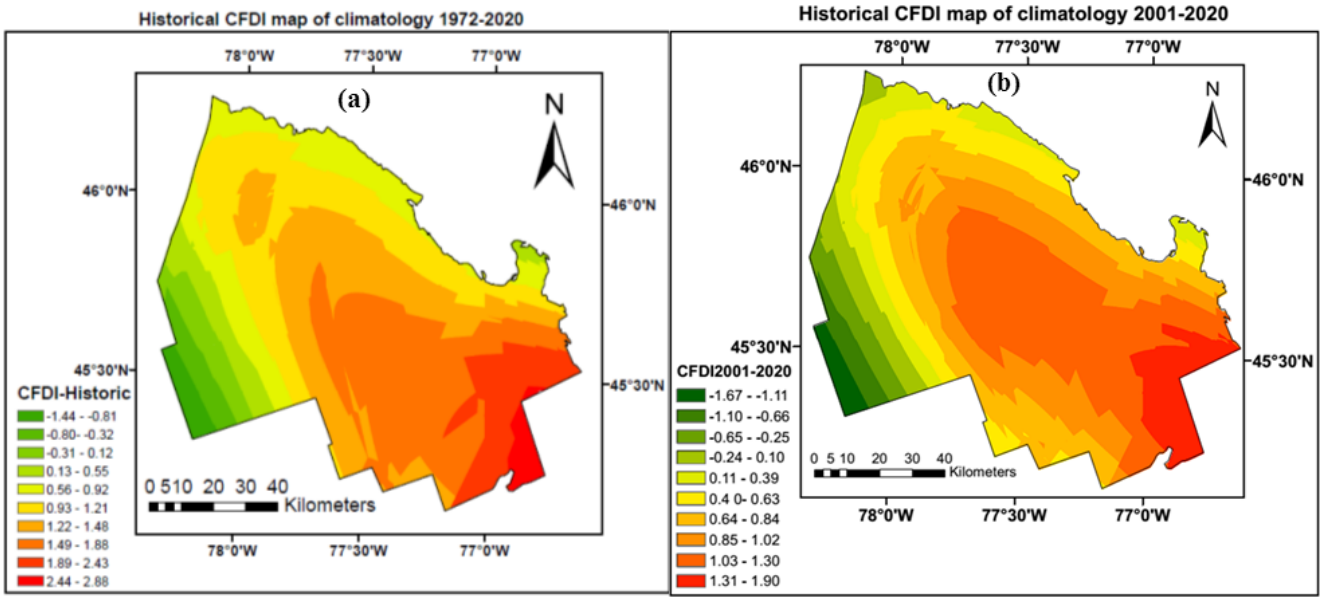


Figure 5. 4 (a-b): a) Historical CFDI gridded maps of climatology created based on the prediction models of (a) 1972–2020 and b) 2001–2020

Based on the differences between historical (1972–2020) (Figure 5.4a) and projected (2050) (Figure 5.3a) CFDI maps, a future forest cover change map was produced (Figure 5.5).

Change Map for 2050 = $CFDI_{\mathbf{Historic}} - CFDI_{\mathbf{2050}}$

The areas of forest cover most vulnerable to climate change were ascertained and highlighted with a color ramp (Figure 5.5). Based on the change map of CFDI 2050 (forest cover difference map), the projected results indicated forest cover reductions and that the southeastern region should be considered extremely vulnerable.

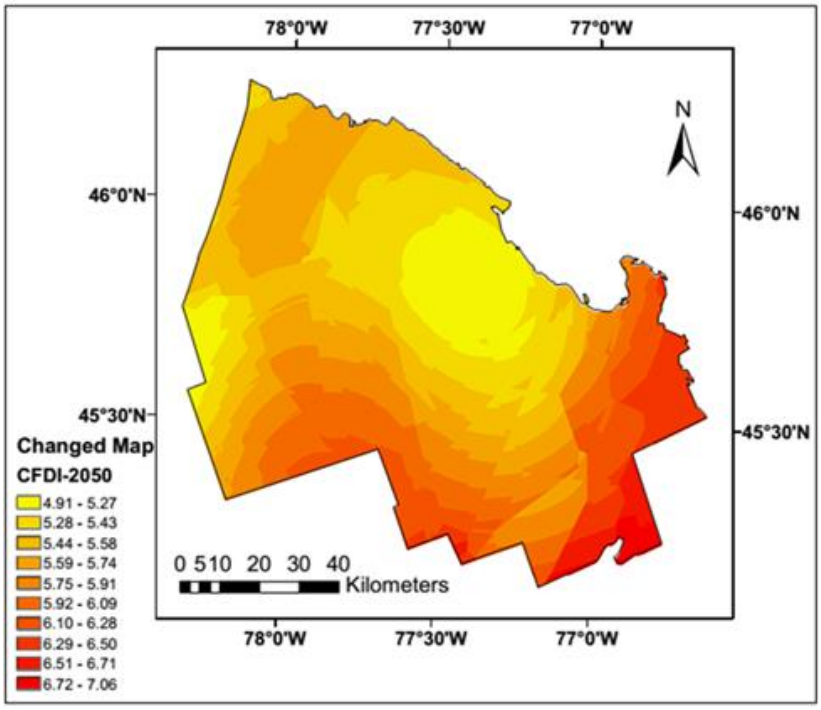


Figure 5. 5: Change map of CFDI 2050

5.4 Discussion

The objective of this study was the development of a set of procedures to ascertain the significant statistical relationships between historic and projected (2050) climate variables and a range of FDIs. The relationships are based on an empirical projected MLR model for the CFDI that could assess the effects of climate change (through the integration of field data, remote sensing, and climate projections) on forest status in an ecologically important study area in Ontario, by using climate variables to project conditions for 2050 under the RCP8.5 scenario.

Daily grids of mean temperature, precipitation, ETo, as well as MA (MA_1 and MA_2) at 30m spatial resolution have been produced for the study area for 1971–2010. This dataset is a complete, multi-decadal, spatiotemporally comprehensive database suitable for regional climate analysis, risk assessment, ecosystem modelling, as well as planning and management purposes. Its production was achieved with a much higher station density, greater spatiotemporal resolution, and longer data record than was possible in many previous studies with gridded databases (McKenney et al., 2006; Thornton et al., 1997).

From the perspective of climate change, all human activities that act to decrease the present carbon stock (carbon carrying capacity) in natural forests need to be included as the primary drivers and pressures on forest degradation (Mackey et al., 2008). Climate change is already affecting the study area, including Algonquin Provincial Park. Earlier ice-out dates occurring on lakes and a general warming trend in the landscape that surround the park over the last 100 years are clear proof of the effects of local climate change. Area trends are a reflection of the rise in daily average temperature during the periods from 1971–2000, 1981–2010, and 2011–2020 of 4.9^0C, 5.1^0C, and 5.2^0C, respectively. However, the most important effects of climatic change in Ontario forests are likely to relate to forest disturbance regimes. In this study, alterations in forest cover (forest degradation) were monitored from 1972–2020 with a computed CFDI resulting from multivariate integration (from seven FDIs) via the image fusion technique (PCA). The

findings also indicated what climate variables (P, MA_1, and MA_2) had the strongest correlation (-0.95, -0.91, and -0.93, respectively) with CFDI. This strong relationship allowed for computing an MLR model with a high coefficient of determination, R^2 (0.93), and low RMSE (0.28). This model was used for projecting the composite CFDI into the future (2050). A comparison of the historical average of 50 years (1971–2020) of climate normals (1971–2010), including average daily climate data (2011–2020), indicated that 2050 will experience daily average temperature and evapotranspiration increases (3.0°C, 0.2 mm, respectively). In addition, there will be a decrease in daily average precipitation (-0.3 mm) and the projected daily average moisture deficits MA_1 and MA_2, will be 0.078 mm and 0.103 mm, respectively. It is also projected that the yearly decrease in precipitation will be 109.5 mm, the ETo increase will be 73.5 mm, and the moisture deficits will be 28.47 mm (MA_1) and 37.60 mm (MA_2). As a result, there will be increased forest degradation as determined by the FDIs.

5.5 Conclusions

In this study, forest cover changes under different climate change scenarios were evaluated based on both the (projected) regression models and projected climate variables. Next, historical and present forest cover maps were produced, and future forest cover maps were also created under the RCP8.5 climate scenario, to identify areas that are more vulnerable in terms of their forest cover. Finally, the difference maps of historical (1971–2020) and projected (2050) forest cover were generated to discover potential changes in forest cover and identify the vulnerable areas in the future.

This study focused on examining the dynamic changes in the forested Algonquin Park study area from 1972–2020 and the generation of predictions of possible changes to 2050. The impacts of changes in temperature (T), precipitation (P), evapotranspiration (ETo), and moisture availability (estimated from two regimes: MA_1 and MA_2) on forest vegetation condition were studied based the effects of temperature and water deficit. To understand the nature of the quantitative relationships between climate variables and

FDIs, MLR analysis was employed to predict future dynamic changes in forest cover. These quantitative relationships were used to compute forest degradation projection maps for the study area. The results furnished evidence that forest cover will likely experience losses (mostly in the south and southeastern portions of the study area) and large fluctuations in the future, putting the area's ecosystem at risk.

Finally, GCM outputs projected a drier climate for the study area. This is a fine example that methodologies for local management on the ground should rely more on GCM projections, which are supposed to act as indicators of future climate change on a large and continental scale. Although extant trends may or may not continue into the future at the same rate, they form the most rational foundation for the development of adaptation strategies, which means they should be employed in combination with GCM projection. RCM output is currently being employed to develop a new generation of climate change scenarios, providing unparalleled opportunities to reach more detailed projections about the future climate. Since they provide both better spatial and temporal resolution, it will be possible to determine the effects of climate change on specific areas. Management changes should be based on models, but adaptation strategies should only be implemented on the ground when the observed trend confirms the projections.

References

Allen, R.G., Pereira, L.S., Raes, D., and Smith, M. 1998. Crop Evapotranspiration: FAO Irrigation and Drainage Paper 56. Rome, Italy.

Amiro, B.D., MacPherson, J.I., Desjardins, R.L., Chen, J.M. and Liu, J. 2003. Post-fire carbon dioxide fluxes in the western Canadian boreal forest: Evidence from towers, aircraft and remote sensing. Agriculture and Forest Meteorology, 115, pp.91–107.

Bárdossy, A., and Plate, E. J. 1992. Space-time model for daily rainfall using atmospheric circulation patterns. Water Resources Research, 28, pp.1247–1259.

Benestad, R.E. 2001. A comparison between two empirical downscaling strategies. International Journal of Climatology, 21(13), pp.1645–1668.

Bronstert, A., Kolokotronis, V., Schwandt, D., and Straub, H. 2007. Comparison and evaluation of regional climate scenarios for hydrological impact analysis: General scheme and application example. International Journal of Climatology, 27, pp.1579–1594.

Browne, S.A., and Hunt L.M. 2007. Climate change and nature-based tourism, outdoor recreation, and forestry in Ontario: Potential effects and adaptation strategies, climate change research report. Ontario Ministry of Natural Resources, Toronto.

Bürger, G., Murdock, T.Q., Werner, A.T., Sobie, S.R., and Cannon, A.J. 2012. Downscaling extremes—An intercomparison of multiple statistical methods for present climate. Journal of Climate, 25, pp.4366–4388.

Canadian Centre for Climate Services: https://climatedata.ca

Chen, J., Brissette, F.P., and Leconte, R. 2014. Assessing regression-based statistical approaches for downscaling precipitation over North America. Hydrological Processes, 28(9), pp.3482–3504 DOI: 10.1002/hyp.9889.

Chen, T., Bao, A., Jiapaer, G., Guo, H., Zheng, G., Jiang, L., Chang, C., and Tuerhanjiang, L. 2019b. Disentangling the relative impacts of climate change and human activities on arid and semiarid grasslands in Central Asia during 1982–2015. Journal of Science of the Total Environment, 653, pp.1311–1325.

Cohen, W. B., Yang, Z., Stehman, S. V., Schroeder, T. A., Bell, D. M., Masek, J. G., Hung, C., and Meigs, G. W. 2016. Forest disturbance across the conterminous United States from 1985–2012: The emerging dominance of forest decline. Forest Ecology and Management, 360, pp.242–252.

Cook, J., Oreskes, N., Doran, P.T., Anderegg, W.R.L., Verheggen, B., Maibach, E.W., Carlton, J.S., Lewandowsky, S., Skuce, A.G., Green, S.A., Nuccitelli, D., Jacobs, P., Richardson, M., Winkler, B., Painting, R., and Rice, K. 2016. Consensus on consensus: A synthesis of consensus estimates on human caused global warming. Environmental Research Letters, 11(4), pp.048002.

Dakhlaoui, H., Ruelland, D., and Tramblay, Y. 2019. A bootstrap-based differential split-sample test to assess the transferability of conceptual rainfall-runoff models under past and future climate variability. Journal of Hydrology, 575, pp.470–486.

Duinker, P.N. 2008. Global climate change. Sustainable forest management network.

Eastaugh, C. 2008. Adaptations of forests to climate change: A multidisciplinary review. IUFRO Secretariat, Vienna.

Environment and Climate Change Canada: https://climate.weather.gc.ca/climate_normals/index_e.html

Environment and Natural Resources Canada. 2018. Canadian Climate Normals 1981–2020. Station Data. Government of Canada, Environment and Natural Resources, Ottawa, Ontario. http://climate.weather.gc.ca/climate_normals.

Environment and Natural Resources Canada: https://climate.weather.gc.ca/historical_data/search_historic_data_e.html

Flannigan, M.D., Amiro, B.D., Logan, K.A., Stocks, B.J., and Wotton, B.M. 2005a. Forest fires and climate change in the 21st century. Mitig Adapt Strategies Glob Change, 11, pp.847–859.

Flannigan, M.D., Logan, K.A., Amiro, B.D., Skinner, W.R., and Stocks, B.J. 2005b. Future area burned in Canada. Climate Change, 72, pp.1–16.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., et al. 2013. Evaluation of climate models. In: Stocker TF et al (eds) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp. 741–866.

Food and Agriculture Organization of the United Nations (FAO). 1978. Report on the Agro-ecological Zones Project. Vol. I. Methodology and Results for Africa. World Soil Resources Report 48/1. FAO, Rome.

Food and Agriculture Organization of the United Nations (FAO). 2012. The ETo Calculator Reference Manual (Version 3.2). Irrigation and Drainage Paper No 56: "Crop Evapotranspiration." Land and Water Division, FAO, Rome, Italy.

Fowler, H. J., Blenkinsop, S., and Tebaldi, C. 2007. Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. International Journal of Climatology, 27(12), pp.1547–1578.

Fowler, H.J., Blenkinsop, S., and Tebaldi, C. 2007. Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. International Journal of Climatology, 27, pp.1547–1578.

Fowler, H.J., Kilsby, C.G., O'Connell, P.E., and Burton, A. 2005. A weather-type conditioned multi-site stochastic rainfall model for the generation of scenarios of climatic variability and change. Journal of Hydrology, 308, pp.50–66.

Frei, C., Christensen, J.H., Déqué, M., Jacob, D., Jones, R.G., and Vidale, P.L. 2003. Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. J. Geophys. Res., 108, pp.1–19.

Frei, C., Schöll, R., Fukutome, S., Schmidli, J., and Vidale, P.L. 2006. Future change of precipitation extremes in Europe: An intercomparison of scenarios from regional climate models. Journal of Geophysical Research, 111, pp.D06105.

Fuglem, P. 1990. Prescribed Burning in British Columbia. B.C. Forest Service.

González-Zeas, D., Erazo, B., Lloret, P., De Bièvre, B., Steinschneider, S., and Dangles, O. 2019a. Linking global climate change to local water availability: Limitations and prospects for a tropical mountain watershed. Science of the Total Environment, 650, pp.2577–2586.

Gu, Z.J., Duan, X.W., Shi, Y.D., Li, Y., and Pan, X. 2018. Spatiotemporal variation in vegetation coverage and its response to climatic factors in the Red River Basin. China. Ecol. Indic., 93, pp.54–64.

Gudmundsson, L., Bremnes, J.B., Haugen, J.E., and Engen-Skaugen, T. 2012. Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods, Hydrol. Earth Syst. Sci., 16, pp.3383–3390, https://doi.org/10.5194/hess-16-3383-2012.

Gullett, D.W., and W.R. Skinner. 1992. The State of Canada's Climate: Temperature Change in Canada 1895-1991. State of the Environment Report No. 92-2. Minister of Supply and Services Canada, pp.36.

Gutiérrez, J. M., San-Martín, D., Brands, S., Manzanas, R. and Herrera, S. 2013. Reassessing statistical downscaling techniques for their robust application under climate change conditions. Journal of Climate, 26(1), pp.171–188.

Häder, D., and Barnes, P. 2019. Comparing the impacts of climate change on the responses and linkages between terrestrial and aquatic ecosystems. Journal of Science of the Total Environment, 682, pp.239–246.

Hewitson, B.C., and Crane, R.G. 1996. Climate downscaling: Techniques and application. Climate Research, 7, pp.85–95

Huete, A., 2016. Ecology vegetation's responses to climate variability. Nature, 531, pp.181–182.

Hughes, J.P., and Guttorp, P. 1994. A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. Water Resources Research, 30(5), pp.1535–1546 DOI: 10.1029/93WR02983.

IPCC, 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the IPCC, Cambridge University Press, Cambridge, United Kingdom and New York, USA.

IPCC. 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA.

IPCC. 2014a. Summary for Policymakers. Contributions of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Core Writing Team: R.K. Pachauri and L.A. Meyer (eds). IPCC, Geneva.

IPCC. 2014b. Synthesis Report. Contributions of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Core Writing Team: R.K. Pachauri and L.A. Meyer (eds). IPCC, Geneva.

IPCC. 2022. Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Pörtner, H.-O., Roberts, D. C., Tignor, M., Poloczanska, E. S., Mintenbeck, K., Alegría, A., Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., and Rama, B. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O.B., Bouwer, L.M., Braun, A., Colette, A., Déqué, M., and Georgievski, G. 2014. EURO-CORDEX: New high-resolution climate change projections for European impact research. Regional Environmental Change, 14, pp.563–578.

Jiapaer, G., Liang, S.L., Yi, Q.X., and Liu, J.P. 2015. Vegetation dynamics and responses to recent climate change in Xinjiang using leaf area index as an indicator. Ecol. Indic., 58, pp.64–76.

Jonson, E.A., and Miyanishi K. 2007. Plant Disturbance Ecology. Academic Press, London.

Kennedy, R.E., Yang, Z., and Cohen, W.B. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. Remote Sen. of Envi., 114, pp.2897−2910.

King, L.M., Mcleod, I.A., and Simonovic, S.P. 2015. Improved weather generator algorithm for multisite simulation of precipitation and temperature. Journal of the American Water Resources Association, 7, pp.1–16.

Kuo, C., Du, R.H., Lin, C.Y., and Yu, P.S. 2005. Using Penman-Monteith method to estimate potential evapotranspiration in Taiwan by using AVHRR and MODIS satellites remote sensing data. In The 26th Asian Conference on Remote Sensing, Hanoi, Vietnam, NSC-91-2211-E-006-038.

Lawal, Sh., Lennard, Ch., and Hewitson, B. 2019. Response of southern African vegetation to climate change at 1.5 and 2.0° global warming above the pre-industrial level. Climate Services Journal, pp.100134.

Lawson, B. D. 1990. "Where There Is Fire, There's Smoke: A Global View of B.C.'s Prescribed Burning," Presentation at Panel Discussion on Smoke Management, Southern Interior Fire Management Committee, Cranbrook, B.C., Canada.

Leung, L. R., Qian, Y., and Bian, X. 2003. Hydroclimate of the western United States based on observations and regional climate. Journal of Climate, 16(12), pp.1892–1911.

Liang X.Z., Kunkel K.E., Meehl, G.A., Jones, R.G., and Wang, J.X.L. 2008. Regional climate models downscaling analysis of general circulation models present climate biases propagation into future change projections. Geophysical Research Letters, 35(8), pp.1–5 DOI: 10.1029/2007GL032849.

Lutz, J.A., van Wagtendonk, J.W., Thode, A.E., Miller, J.D., and Franklin, J.F., 2009. Climate, lightning ignitions, and fire severity in Yosemite National Park, California, USA. Int. J. Wildl. Fire, 18, pp.765–774.

Ma, Q., Zhu, L., Wang, J., Ju, J., Wang, Y., Lu, X., Kasper, Th., and Haberzettl, T. 2020. Late Holocene vegetation responses to climate change and human impact on the central Tibetan Plateau. Journal of Science of the Total Environment, 708, pp.135370.

Mackey, B.G., Keith, H., Berry, S., Lindenmayer, D.B. 2008. Green carbon: The role of natural forests in carbon storage. Part 1, A green carbon account of Australia's southeastern Eucalypt forest, and policy implications. ANU E Press, Canberra, Australia.

MacLean, I.M.D., Suggitt, A.J., Wilson, R.J., Duffy, J.P., and Bennie, J.J. 2017. Fine-scale climate change: Modelling spatial variation in biologically meaningful rates of warming. Global Change Biology, 23, pp.256–268.

Maurer, E.P., Hidalgo, H.G., Das, T., Dettinger, M.D., and Cayan, D.R. 2010. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrol. Earth Syst. Sci., 14, pp.1125–1138, https://doi.org/10.5194/hess-14-1125-2010.

McDermid, J., Fera, S., and Hogg, A. 2015. Climate change projections for Ontario: An updated synthesis for policymakers and planners. Ontario Ministry of Natural Resources and Forestry, Science and Research Branch, Peterborough, ON. Climate Change Research Report CCRR-44.

Mearns, L.O., Giorgi, F., Whetton, P., Pabon, D., Hulme, M. and Lal, M. 2003. Guidelines for use of climate scenarios developed from regional climate model experiments. IPCC Task Group on Scenarios for Climate Impact Assessment (TGCIA). Retrieved from the Data Distribution Centre of the Intergovernmental Panel on Climate Change.

Meehl, G.A., Boer, G.J., Covey, C., Latif, M., and Stouffer, R.J. (2000). The coupled model intercomparison project (CMIP). Bulletin of the American Meteorological Society, 81, pp.313–318.

Mehrotra, R., and Sharma, A. 2005. A nonparametric nonhomogeneous hidden Markov model for downscaling of multisite daily rainfall occurrences. Journal of Geophysical Research, 110(16), pp.1–13.

Mo, K., Chen, Q., Chen, C., Zhang, J., Wang, L., Bao, Z. 2019. Spatiotemporal variation of correlation between vegetation cover and precipitation in an arid mountain-oasis river basin in northwest China. J. Hydrol., 574, pp.138–147.

Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., Van Vuuren, D.P., Meehl, G.A. et al. 2010. The next generation of scenarios for climate change research and assessment. Nature, 463(7282), pp.747–756.

Nadeau, C.P., Urban, M.C., and Bridle, J.R. 2017. Coarse climate change projections for species living in a fine-scaled world. Global Change Biology, 23, pp.12–24.

NASA (National Aeronautics and Space Administration). 2018. Global land-ocean temperature index. National Aeronautics and Space Administration, Goddard Institute of Space Studies. California Institute of Technology, Passadena, CA, USA. https://climate.nasa.gov/vital-signs/global-temperature/.

Ontario Ministry of Natural Resources. 2008b. The boreal forest. Ontario Ministry of Natural Resources.

Onyutha, C., Tabari, H., Rutkowska, A., Nyeko-Ogiramoi, P., and Willems, P. 2016. Comparison of different statistical downscaling methods for climate change rainfall projections over the Lake Victoria basin considering CMIP3 and CMIP5. Journal of Hydro-environment Research, 12, pp.31–45.

Parlange, M.B., and Katz, R.W. 2000. An extended version of the Richardson model for simulating daily weather variables. Journal of Applied Meteorology, 39(5), pp.610–622.

Pfizenmayer, A., and von Storch, H. 2001. Anthropogenic climate change shown by local wave conditions in the North Sea. Climate Research, 19(1), pp.15–23.

Qian, B., Gameda, S., de Jong, R., Falloon, P., and Gornall, J. 2010. Comparing scenarios of Canadian daily climate extremes derived using a weather generator. Climate Research, 41, pp.131–149.

Ren, S.L., Li, Y.T., and Peichl, M., 2020. Diverse effects of climate at different times on grassland phenology in mid-latitude of the Northern Hemisphere. Ecol. Indic., 113, pp.106260.

Ridgway, M., Smith, D., and Middel, T. 2018. Climate warming projections for Algonquin Provincial Park. Ontario Ministry of Natural Resources and Forestry, Science and Research Branch, Peterborough, ON. Science and Research Information Report IR-14.

Ruiz-Benito, P., Vacchiano, G., Lines, E., Reyer, C., Morin, X., Hartig, F., Makela, A., Yousefpour, R., Chaves, J., Palacios-Orueta, A., Benito-Garzon, M., Molino, C., Camarero, J., Jump, A., Kattge, J., Lehtonen, A., Ibrom, A., and Zavala, M. 2020. Available and missing data to model impact of climate change on European forests. Ecological Modelling Journal, 416, pp.1–15.

Rummukainen, M. 2010. State-of-the-art with regional climate models. Wiley Interdisciplinary Reviews: Climate Change, 1(1), pp.82–96.

Schindler, D.W. 1998. A dim future for Boreal waters and landscapes. Bioscience, 48, pp.157 –164.

Schirpke, U., Kohler, M., Leitinger, G., Fontana, V., Tasser, E., Tappeiner, U. 2017. Future impacts of changing land-use and climate on ecosystem services of mountain grassland and their resilience. Ecosystem Service Journal, 26, 79–94.

Shi, H.Y., and Chen, J. 2018. Characteristics of climate change and its relationship with land use/cover change in Yunnan Province. China. Int. J. Climatol., 38(5), pp.2520–2537.

Shi, H.Y., Chen, J., Wang, K.Y., and Niu, J. 2018. A new method and a new index for identifying socioeconomic drought events under climate change: A case study of the East River basin in China. Sci. Total Environ., pp.616–617; 363–375.

Shi, H.Y., Li, T.J., and Wang, G.Q. 2017. Temporal and spatial variations of potential evaporation and the driving mechanism over Tibet during 1961–2001. Hydrol. Sci. J., 62(9), pp.1469–1482.

Soja, A.J. Tchebakova, N.M., French, N.H.F., Flannigan, M.D., Shugart, H.H., Stocks, B.J., Sukhinin, A.I., Parfenova, E.I., Chapin, F.S. III, and Stackhouse, P.W. Jr. 2007. Climate-induced boreal forest change: Predictions versus current observations. Global Planet Change, 56, pp.274–296.

Solman, S.A. 2013. Regional climate modeling over South America: A review. Advances in Meteorology, pp.1–13.

Srivastav, R.K., and Simonovic, S.P. 2014. Multi-site, multivariate weather generator using maximum entropy bootstrap. Climate Dynamics, 44, pp.3431–3448.

Stern, N. 2007. The Economics of Climate Change: The Stern Review. Cambridge, Cambridge University Press.

Sun, P., Yu, Z., Liu, S., Wei, X., Wang, J., Zegre, N., and Liu, N., 2012. Climate change, growing season water deficit and vegetation activity along the north-south transect of eastern China from 1982 through 2006. Hydrol. Earth Syst. Sci., 16(10), pp.3835–3850.

Sun, Y., Shan, M., Pei, X., Zhang, X., and Yang, Y. 2020. Assessment of the impacts of climate change and human activities on vegetation cover change in the Haihe River basin, China. Journal of Physics and Chemistry of the Earth, Parts A/B/C, pp.102834.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. 2012. An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society, 93, pp.485–498.

Trenberth, K.E. 2011. Changes in precipitation with climate change. Climate Research, 47, pp.123–138.

Turner, M.G. 1989. Landscape ecology: The effect of pattern on process. Annu. Rev. Ecol. Syst., 20, pp.171–197.

Van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., and Lamarque, J. F. 2011. The representative concentration pathways: An overview. Climatic Change, 109, pp.5–31.

Volney, W.J.A., and Hirsch, K.G. 2005. Disturbing forest disturbances. For. Chron., 81(5), pp.662–668.

Volosciuk, C., Maraun, D., Semenov, V.A., and Park, W. 2015. Extreme precipitation in an atmosphere general circulation model: Impact of horizontal and vertical model resolutions. Journal of Climate, 28, pp.1184–1205.

Von Storch, H., Zorita, E., and Cubasch, U. 1993. Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. Journal of Climate, 6, pp.1161–1171.

Wang, B., Xu, G., Li, P., Li, Z.B., Zhang, Y.X., Cheng, Y.T., Lu, J., and Zhang, J.X. 2019. Vegetation dynamics and their relationships with climatic factors in the Qinling Mountains of China. Ecol. Indic., 108, pp.105719.

Westerling, A.L., Turner, M.G., Smithwick, E.A.H., Romme, W.H., and Ryan, M.G. 2011. Continued warming could transform Greater Yellowstone fire regimes by mid-21st century. Proc. Natl. Acad. Sci., 108, pp.13165–13170.

Wilby R.L., Dawson C.W., and Barrow, E.M. 2002. SDSM — A decision support tool for the assessment of regional climate change impacts. Environmental Modelling & Software, 17(2), pp.147–159.

Wilby, R.L. and Wigley, T.M.L. 1997. Downscaling general circulation model output: A review of methods and limitations. Progress in Physical Geography, 21, pp.530–548.

Wilby, R.L., and Wigley, T.M.L. 2000. Precipitation predictors for downscaling: Observed and general circulation model relationships. International Journal of Climatology, 20(6), 641–661.

Wilby, R.L., Charles, S.P., Zorita, E., Timbal, B., Whetton, P., and Mearns, L.O. 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods. Supporting material of the intergovernmental panel on climate change. Prepared on behalf of Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA). Retrieved from the Data Distribution Centre of the Intergovernmental Panel on Climate Change.

Wilks, D. S., and Wilby, R.L. 1999. The weather generation game: A review of stochastic weather models. Progress in Physical Geography, 23(3), pp.329–357.

Wilks, D.S. 2010. Use of stochastic weather generators for precipitation downscaling. Wiley Interdisciplinary Reviews: Climate Change, 1, pp.898–907.

Wood, A.W., Leung, L.R., Sridhar, V., and Lettenmaier, D.P. 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic Change, 62, pp.189–216.

Wulder, M.A. Campbell, C. White, J.C., Flannigan, M.D., and Campbell, I.D. 2007. National circumstances in the international circumboreal community. For Chron, 83(4), pp.539–556.

Yang, Q., Zhang, H., Peng, W., Lan, Y., Luo, S., Shao, J., Chen, D., and Wang, G. 2019. Assessing climate impact on forest cover in areas undergoing substantial land cover change using Landsat imagery. Science of the Total Environment Journal, 659, pp.732–745.

Yao, J., Chen, Y., Zhao, Y., Mao, W., Xu, X., Liu, Y., and Yang, Q. 2018. Response of vegetation NDVI to climatic extremes in the arid region of Central Asia: A case study in Xinjiang. China. Theor. Appl. Climatol., 131, pp.1503–1515.

Zhang, K., Wie, J.Z., Pei, J.Y., Cheng, H., Zhang, X.L., Huang, F.Q., Li, F.M., and Ye, J.S. 2019b. Impacts of climate change and human activities on grassland vegetation variation in the Chinese Loess Plateau. J. Sci. Tot. Environ., 660, pp.236–244.

Zhang, P.P., Cai, Y.P., Yang, W., Yi, Y.J., Yang, Z.F., and Fu, Q. 2019. Multiple spatio-temporal patterns of vegetation coverage and its relationship with climatic factors in a large dam-reservoir-river system. Ecol. Eng., 138, pp.188–199.

Zhang, W., Wang, L.C., Xiang, F.F., Qin, W.M., and Jiang, W.X. 2020. Vegetation dynamics and the relations with climate change at multiple time scales in the Yangtze River and Yellow River Basin, China. Ecol. Indic., 110, pp.105892.

Zhang, Y., Gao, J., Liu, L., Wang, Z., Ding, M., and Yang, X. 2013. NDVI-based vegetation changes and their responses to climate change from 1982 to 2011: A case study in the Koshi River Basin in the middle Himalayas. Ecol. Indic., 108, pp. 139–148.

Zhao, J., Huang, S.Z., Huang, Q., Wang, H., Leng, G.Y., and Fang, W. 2020. Time-lagged response of vegetation dynamics to climatic and teleconnection factors. Catena, 189, pp.104474.

CHAPTER SIX

CONCLUSIONS

This research developed a methodological framework for the continuous assessment, mapping, and monitoring of present and historic forest disturbance and degradation trends, and their future projections under climate change. The framework rests mainly on remote sensing technology, supported by ground measurements, and predictive models. The methodological framework identified and tested procedures for the continuous monitoring of forest disturbance and degradation in a forested area of Ontario.

First, this thesis examined and evaluated the accuracy of conventional and relatively new artificial intelligence algorithms (i.e., machine learning) for land cover classification based on Landsat imagery of Central and Northeastern Ontario, Canada. A map with the highest possible level of accuracy for the study area was created as a result of these efforts to address the lack of an accurate existing land cover map. In the process, four algorithms—two non-parametric machine learning algorithms: support vector machine (SVM) and artificial neural networks, as well as two parametric classification algorithms: maximum likelihood classifier and Mahalanobis distance. The results demonstrated that SVM is the most accurate, robust, consistent, and effective algorithm for Landsat image classification in all investigated metrics, including in terms of imbalanced training data. For SVM, the amount of time and effort invested led to gains in accuracy from 60.86% in version 1 to 98.62% in version 8, with a time investment of 14hrs and 42hrs, respectively.

A suitable set of forest degradation indicators (FDIs) derived from multispectral satellite imagery was identified and computed and then compared to those of international standards. An innovative composite forest degradation indicator (CFDI) emerged from the synthesis of FDIs, which was employed to examine its relationship with climate variables and projected future (2050) forest degradation under climate

change. Long-term trajectory analysis of the FDIs, including the CFDI, proved useful in the development of disturbance and degradation maps for 1972–2020.

An innovative component of this research was the fusion of Landsat and MODIS (in spite of their differences in spatial and temporal resolutions) computed FDIs (NDVI, NPP, NBR, EVI, TCB, TCG, TCW) into a combined and more explanatory CFDI that is the most informative indicator of combined forest degradation, for a combined map for 1972–2020. The disturbance maps were found to be effective in improving our understanding of the spatio-temporal land cover change trajectories from 1972–2020. Overall, the innovative CFDI (1972–2020) was the most informative as the combination of three layers of CFDI (1972–1981, 1984–1998, and 2001–2020), and the explained variability of the combined CFDI was 95.32%.

The LandTrendr algorithm was able to clearly identify patterns of forest disturbance, including the duration and progress of disturbance and recovery. Therefore, this study revealed how critical it is to synergistically use archived Landsat and MODIS images with trajectory analysis to continuously monitor land cover change.

A geostatistical analysis of relevant climate variables related to FDIs computed from available multispectral image stacks found relationships that were useful in projecting FDIs through multiple linear regression (MLR) models, which were used to examine the impact of climate change on forest cover, and to provide a forest vegetation cover degradation map under the projected climate scenario of 2050. The relationships between remote sensing indicators and climate variables led to the creation of spatially-explicit long-term historical forest degradation maps.

The results for the period from 1972–2020 indicated climate variables P, MA_1, and MA_2 had the strongest correlation (-0.95, -0.91, and -0.93, respectively) with CFDI, which enabled for computing of a model with a high coefficient of determination, R^2 (0.93), and low RMSE (0.28). Additionally, findings for

2001–2020 showed climate variables P, MA_1, and MA_2 had the most significant correlation with CFDI, NDVI, and NPP. This strong relationship allowed for the establishment of future models with a high coefficient of determination, R^2 (0.83, 0.80, 0.92, respectively), and low RMSE (0.39, 0.38, 0.26, respectively).

The Ontario study area forests displayed evidence of moisture stress, particularly in the summer months (June–August) from 1971–2010. The impacts of changes in temperature, precipitation, evapotranspiration, and moisture availability on forest vegetation conditions showed evidence that forest cover is likely to suffer major losses and fluctuations in the future as related to climate variables (especially moisture availability), which will put the area's entire ecosystem at risk. Thus, this method of monitoring forest degradation through the integration of remote sensing tools could assist and certainly improve climate change risk management, in particular by providing data on climate effects on vegetation. Potential changes in forest cover and vulnerable areas in the future, according to the projected climate under the representative concentration pathway 8.5 scenario were mapped. Based on the differences between the historical (past and present) and future (year 2050) CFDI map, the future forest cover map predicts forest cover reductions, mostly in the south and southeast of the study area.

Projecting the CFDI into the future is a useful tool for visualizing the future of forests in the area. Based on 50 years of historical data, it is possible to predict that 2050 will experience daily average temperature and evapotranspiration increases (3.0°C, 0.2 mm, respectively). There will also be a decrease in daily average precipitation (0.3 mm), and the projected daily average moisture deficits, MA_1 and MA_2, will be 0.078 mm and 0.103 mm, respectively. It is also projected that the yearly decrease in precipitation will be 109.5 mm, the evapotranspiration increase will be 73.0 mm, and the moisture deficits will be 28.47 mm (MA_1) and 37.60 mm (MA_2). Additionally, study area temperature trends are a reflection of the rise in daily average temperature, which was 4.9^0C, 5.1^0C, and 5.2^0C for the years from 1971–2000, 1981–2010,

and 2011–2020, respectively. All the factors detailed here are likely to lead to increased forest degradation as determined by the FDIs.

This study developed a fast and effective approach for monitoring forest disturbance and degradation from multiscale and multi-temporal remote sensing data. Exploiting the "big data" stacks from multispectral sensors and combining spatial and temporal resolutions has significant potential in monitoring regional scale forest degradation. The FDIs used in this study could effectively detect long-term forest disturbance and degradation. The results of this thesis showed that Landsat and MODIS satellite platforms can be combined and used to derive FDIs, which were effective in the identification and mapping of forest disturbance and degradation over the long-term. The analysis of the spectral response of vegetation with both remote and proximal sensing could prove to be an efficient and rapid way to examine and define forest health. However, more effective data analysis procedures should be developed, especially given the ongoing increase in the amount of data furnished by numerous sensors.

This research also bridged current knowledge related to the spatial distribution, temporal trends, and patterns in the relationship between climate variables and FDIs, in order to provide forest degradation prediction maps for the study area. Since the negative impacts of losses and large fluctuations in vegetation in the future may lead to ecosystem vulnerability, this method of monitoring forest degradation through the integration of tools could assist and certainly improve climate change risk management and planning, in particular by providing data on climate impacts on forest vegetation. The consequences of projected forest degradation with climate change in the region of study relative to greenhouse gas concentrations and fluxes were not addressed in this study. Therefore, future work should address this gap to aid in the development of appropriate climate change mitigation and adaptation strategies for forest ecosystem and biodiversity management for the region.

General circulation model projections should be employed in combination with regional climate model outputs to develop climate change scenarios providing more detailed projections about the future climate with better spatial and temporal resolution. Management changes should be based on models, but adaptation strategies should only be implemented on the ground when the observed trend confirms the projections.